# CHARACTERIZATION OF ADENO-ASSOCIATED VIRUS 2 SITE-SPECIFIC INTEGRATION

By

SHYAM DAYA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2009

To Almighty and my Parents

ACKNOWLEDGMENTS

First of all, I would like to thank Dr. Kenneth Berns for giving me the opportunity to do research in his laboratory, for his faith in my abilities, and for his complete support throughout my graduate education.

Secondly, I want to express my gratitude to all the members of my graduate committee: Dr. Mavis Agbandje-McKenna, Dr. Arun Srivastava, and Dr. Bradley Fletcher. I couldn't have selected a better group of professors.

Thirdly, I would also like to thank my laboratory colleagues, Nenita Cortez and Chun Zhang. I would not be able to graduate on time without Nenita's help. Chun Zhang had been a good source of early guidance when I joined the laboratory.

Last but not least, I have received great support from my family and the Almighty. I thank my family for their love and thank the Almighty as a constant source of support and inspiration.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CHARACTERIZATION OF ADENO-ASSOCIATED VIRUS 2 SITE SPECIFIC INTEGRATION

By

Shyam Daya

May 2009

Chair: Kenneth I. Berns
Major: Medical Sciences

Adeno-associated virus (AAV) is the only mammalian DNA virus known to be capable of establishing latency by integration into a specific site, called AAVS1, on Chromosome 19q13.4. The AAV *cis* and *trans* requirements for this process have been identified, yet, a complete picture has not emerged with regards to the precise mechanism. Specifically, the host protein requirements are not known, as well as how they could be interacting with the AAV genome.

This study has focused on two specific research objectives. The first objective was to determine if the AAV Rep protein, shown to be required for targeting site-specific integration, can minimize AAV random integration. The second objective was to study cellular repair proteins for an effect on AAV site specific integration.

My data has indicated that even in the presence of Rep, random integrations do take place, but, the limited number of identified AAV-cellular junctions precludes any definitive statements as to whether Rep can minimize random integration. In addition, I presented direct evidence that a protein called DNA-dependent protein kinase catalytic subunit (DNAPKcs) inhibits stable site-specific integration of single-stranded AAV Rep positive vectors. Moreover, the presence or absence of DNAPKcs did not affect the specific integration of self-complementary AAV Rep

positive vectors, which also displayed more random integration. In addition, cellular repair proteins ligase I and ligase IV are not needed for AAV-AAVS1 junction formation, but the absence of ligase IV greatly reduced the frequency of AAV site-specific integration.

These findings contribute to a better understanding of the AAV site-specific integration process, suggesting that components of the non-homologous end joining pathway can modulate AAV site-specific integration. By systematically studying the different cellular repair proteins for an effect on specific integration, future work should be able to elucidate the mechanism. The outcome of such advancements would lead to an increased emphasis for the further development of AAV as a safe and efficient integrating vector for future gene therapy studies.

# CHAPTER 1
# INTRODUCTION

## Gene Therapy

Gene Therapy proposes the simple concept that the introduction of specific genetic material (DNA or RNA), can compensate for or modulate the expression of the defective endogenous gene causing the disease, resulting in a therapeutic outcome. This complex field requires that researchers achieve a comprehensive understanding of the genetic disorder to be treated with regards to the nature of the genetic defect, types of tissue or cells that need to be targeted for therapeutic expression, type of expression required (regulated or constitutive), and duration of expression (short term or long term). Early research in this field has focused largely on developing delivery approaches. As of now, the emphasis has been on improving the delivery approaches to target expression to specific cells and tissues, and understanding the host immune response to gene transfer.

In general, two approaches exist that aim to introduce genetic material into target cells and tissues: non-viral and viral. In non-viral gene therapy, genetic material (i.e. DNA or DNA-liposome complexes) is introduced directly into target tissues or cells. In viral-mediated gene therapy, human and non-human primate viruses are used to deliver genetic material. Viral gene therapy has gained popularity over non-viral gene therapy because viruses possess the natural ability to infect cells and tissues. Moreover, the many available viruses, allows for both DNA and RNA to be delivered. Some of the most commonly used RNA-based viruses are retrovirus and lentivirus. Herpes simplex virus, adenovirus, and adeno-associated virus (AAV) are commonly used DNA-based viruses.

There are two strategies for both these gene therapy approaches: 1) *ex vivo*, where target cells are transduced outside the body and transplanted back into the body, and 2) *in vivo*, where

target cells are directly transduced within the body. Non-viral gene transfer suffers from very low *in vivo* efficacy, and has benefited greatly from the *ex vivo* approach for gene therapy. In contrast, viral gene therapy can be effectively used in both *in vivo* and *ex vivo* strategies. Therefore, gene therapy applications have now focused largely on developing viruses for gene transfer.

In 1990, scientists successfully used a retrovirus to treat a form of severe combined immunodeficiency disorder, adenosine deaminase deficiency (ADA-SCID), in a patient. An *ex vivo* strategy was used to target gene expression of the ADA gene in the bone marrow cells of the patient. This outcome brought great popularity to the field. However, this kind of success was not seen in several other gene therapy clinical trials. For example, in 1999, a patient by the name of Jesse Gelsinger, died from a gene therapy trial which was aimed at treating ornithine transcabamylase (OTC) deficiency using adenovirus as a delivery vehicle (Raper et al., 2002). His death occurred as a result of the immune response towards the adenovirus capsid (Raper et al, 2003). It remains unknown why Gelsinger suffered these adverse effects, because a second patient tolerated the same treatment dose. Nevertheless, this clinical outcome served as major set-back to the gene therapy field.

Moreover, results from another clinical trial demonstrated that gene therapy was still in its infancy. Nine of the ten treated patients were successfully treated for the fatal X-linked severe combined immunodeficiency disorder (X-SCID) using a retroviral vector carrying the gene encoding the gamma-c chain cytokine receptor; (Cavazzana-Calvo et al., 2000); however, a follow up revealed that 4 of the 9 treated patients had later developed acute lymphoblastic leukemia (Hacein-Bey-Abina et al., 2008).

These examples briefly present the turbulent history for the gene therapy field and highlight the challenges faced by gene therapy researchers. In order to develop safer and therapeutic gene delivery, researchers are now studying both non-viral and viral approaches in depth to understand their limitations, their biological properties, and their interactions with the host immune system.

## Gene Therapy Using AAV Vectors

The development and use of AAV for gene therapy has been promising to the gene therapy field. AAV is unique when compared to other viruses in that it has not been shown to cause disease. Moreover, it is a naturally defective virus that requires a helper virus co-infection for productive infection, and it can infect both non-dividing and dividing cells, making it useful for both *in vitro* and *ex vivo* gene transfer strategies.

AAV-based viral vectors are being tested for treatment of cystic fibrosis, hemophilia, alpha-1-antitrypsin deficiency, Canavan's disease, muscular dystrophy, Alzheimers, and Parkinsons to name a few. According to the Journal of Gene Medicine, AAV is being used in 60 ongoing clinical trials, accounting for roughly 4% of all clinical trials. Some very recent phase I clinical trial success has been achieved with AAV vectors aimed at treating Leber's congential amaurosis (LCA) (Bainbridge et al., 2008; Hauswirth et al., 2008; Maguire et al., 2008). In LCA, a defect exists in the expression of RPE65 protein, which is required for production of rhodopsin. Since rhodopsin is needed for photoreceptor function in sight, patients with LCA suffer from poor vision early during childhood that later progresses to blindness. These LCA clinical trials provided evidence for the safety of the delivered AAV viral vector and demonstrated improved visual function in most of the patients that were treated. This outcome represents the first success in the use of AAV as a vector in a human clinical trial, and provides great promise for the use of AAV for other diseases.

Before the LCA trials, a phase I study was initiated to treat Hemophilia B, using AAV to deliver Factor IX into the livers of seven patients (Manno et al., 2006). Therapeutic levels was observed in several patients given the highest doses, but these levels were short lived, dropping to baseline (< 1 % serum level of Factor IX) by 10 weeks. One patient, in particular, had an asymptomatic rise in liver transaminase levels, suggestive of liver damage, which returned to normal levels by 9 weeks. The coincident decline of Factor IX levels and the rise of transaminase levels indicated that transduced cells were being subjected to destruction by the immune system. A subsequent study identified that this response was due to cytotoxic T-cell response towards the AAV 2 transduced hepatocytes (Mingozzi et al., 2007).

Taken together, these studies highlight the limited success of AAV in clinical trials, and suggest that immune response towards AAV vectors can present major hurdles to gene transfer. As researchers achieve a deeper understanding of AAV-host interactions in humans, better AAV vectors can be developed that have improved safety and efficacy profiles for treatment of many acquired and inherited diseases.

## Purpose and Scope of This Research

AAV may serve as a successful gene delivery vehicle. However, the advancements in understanding this virus and its development as a viral vector have fallen short in terms of success in clinical trials. A great deal remains unknown about the infection process of AAV and its persistence in different cells, tissues, and organs.

There are many different types (i.e serotypes) of AAV. AAV serotype 2 (AAV 2) is the prototypical serotype which has been primarily used in vector development and in most of the clinical trials. The infection process of AAV 2 has been studied both in cell culture and *in vivo*. AAV 2 has two phases to its life cycle: a lytic phase and a latent phase. The lytic phase is characterized by replication of the AAV genome and the production of newly packaged AAV

particles, in the presence of a helper virus. In the absence of a helper virus, AAV establishes latency. In cell culture, this latent phase is characterized by the integration of the AAV genome into a specific-chromosomal site, termed AAVS1. A great deal is known about the role of the helper virus and cellular elements during lytic infection, however, information is lacking about which cellular proteins are involved in AAV latent infection.

The purpose of my research was to understand more about the latent infection process of AAV with regards to the specificity of integration and the cellular requirements. The AAV replication protein (Rep) has been shown to be necessary for targeting latency site-specifically. Very little is known with regard to the cellular requirements. Many researchers have attempted to identify Rep interacting cellular proteins in hopes of identifying those cellular factors. Recently, Nash et al. identified 188 Rep-interaction proteins, and classified them into many categories based on their role in DNA replication, repair, transcription, and splicing to name a few (Nash et al., 2009). AAV may be using the cellular DNA repair machinery to help it integrate, but the lack of available methodologies to study integration has hindered the identification of the precise proteins. It is anticipated that my work could serve as a framework from which future researchers can then start identifying those cellular factors. Moreover, it should provide an emphasis for the development of AAV as novel, site-specifically integrating vector for gene therapy applications requiring long term expression of a therapeutic gene.

## Outline of Chapters

Chapter two, Literature Review, will describe what is known about the biology of AAV. The first part covers important information about the AAV virus, its genome, and its biphasic life cycle of expression and latency. The second part of the chapter briefly covers how recombinant AAV vectors are created and packaged into AAV particles. Moreover, advancements in AAV vector development are described which should improve the utility of AAV vectors. This chapter

ends with a short section on the fate of recombinant AAV vectors in transduced tissues and cells.

Chapter three, Research Objectives, provides a brief background about AAV site-specific integration to provide a foundation for understanding the details of the research. Chapters four and five cover the first and second aims of my project, respectively. The first specific aim describes the studies that were done to determine the specificity of AAV site-specific integration, looking at the distribution of AAV integration in the presence of the AAV replication protein (Rep), to determine if Rep can minimize random integration events. The second specific aim, describes the studies that were aimed at determining if cellular repair proteins can modulate AAV site-specific integration. Specifically, the role of a non-homologous repair protein, DNA-dependent protein kinase catalytic subunit (DNAPKcs) was investigated in some detail. This protein functions as a crucial element in the cellular repair of DNA via the non-homologous end joining pathway.

CHAPTER 2
LITERATURE REVIEW

## Biological Properties of AAV

### AAV Structure and Genome

Adeno-associated virus is a small, non-enveloped, icosahedral virus (25 nm), which was discovered among adenovirus preparations (Melnick et al., 1965; Atchison et al., 1965; Hoggan et al., 1966). It belongs to the *Parvoviridae* family and the genus *Dependovirus*, because it requires a helper virus for productive infection (Hoggan et al., 1968). AAV packages a single stranded DNA genome, and equal numbers of both positive and negative sense strands are packaged in separate virions with equivalent efficiency (Rose et al., 1969; Mayor et al., 1969).

Thirteen serotypes of AAV (AAV serotype 1 [AAV 1] to AAV 13) have been discovered to date. The prototypical AAV serotype, AAV 2, is the best studied. The structure of AAV 2 has been solved by X-ray crystallography (Xie et al., 2002). Major constituents of the AAV 2 capsid are the many β secondary structures which are arranged into β-barrels. These β barrels are connected by loops of variable lengths, which are exposed on the capsid surface. Interestingly, these loops are not well conserved among the AAV serotypes and the other parvoviruses, accounting for the unique capsid structure and function observed for the different serotypes. The prominent features of the AAV capsid surface include a three-fold spike, a two-fold depression, and a five-fold pore.

In addition, over 100 AAV variants have been isolated from both primate and human tissues, which have been arranged into different clades (Gao et al., 2004), based on their capsid homology to the other AAV types. A comparison of the capsid homology among AAV serotypes 1 through 13 is shown in Table 2-1. AAV serotypes 1 through 9 have been studied in some detail (Zincarelli et al., 2008) and have clinical applications for humans. Interestingly,

AAV serotypes 1 and 6 share significant capsid homology, as do AAV serotypes 3 and 13. Surprisingly, Wu et al. found that a single amino acid difference at residue 531 between AAV 1 and 6 differentially affected heparin binding and liver transduction profiles of these very similar serotypes (Wu et al., 2006). This suggests that functional differences with respect to receptor usage and infection properties of AAV serotypes can be mapped to a single residue or several critical residues, despite high capsid sequence homology. In addition, it implies that all these AAV serotypes can be modified for specific gene therapy applications, conditional upon developing a better understanding of the structure-function and host interactions of these serotypes.

Table 2-1. Percent capsid homology among AAV serotypes 1 through 13.

| | AAV 1 | AAV 2 | AAV 3 | AAV 4 | AAV 5 | AAV 6 | AAV 7 | AAV 8 | AAV 9 | AAV 10 | AAV 11 | AAV 12 | AAV 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AAV1 | 100 | | | | | | | | | | | | |
| AAV2 | 83 | 100 | | | | | | | | | | | |
| AAV3 | 86 | 87 | 100 | | | | | | | | | | |
| AAV4 | 63 | 60 | 62 | 100 | | | | | | | | | |
| AAV5 | 58 | 57 | 58 | 53 | 100 | | | | | | | | |
| AAV6 | 99 | 83 | 87 | 63 | 58 | 100 | | | | | | | |
| AAV7 | 85 | 82 | 85 | 63 | 58 | 85 | 100 | | | | | | |
| AAV8 | 84 | 83 | 86 | 63 | 58 | 84 | 88 | 100 | | | | | |
| AAV9 | 82 | 82 | 83 | 62 | 57 | 82 | 82 | 85 | 100 | | | | |
| AAV10 | 85 | 84 | 85 | 63 | 57 | 85 | 88 | 93 | 86 | 100 | | | |
| AAV11 | 66 | 63 | 63 | 81 | 53 | 66 | 67 | 65 | 63 | 66 | 100 | | |
| AAV12 | 60 | 60 | 60 | 78 | 52 | 60 | 62 | 62 | 60 | 61 | 84 | 100 | |
| AAV13 | 87 | 88 | 94 | 65 | 58 | 87 | 85 | 85 | 84 | 86 | 65 | 60 | 100 |

AAV 2 packages a 4.7 kb genome (Figure 2-1), comprised of two genes, rep and cap, flanked by inverted terminal repeats (ITRs) of 145 nucleotides (Srivastava et al., 1983). The terminal 125 nucleotides of each ITR form a palindrome which folds upon itself via base pairing to create a T-shaped hairpin structure. The ITRs are important *cis* sequences in the biology of AAV, being required for DNA replication, packaging of the AAV genome, transcription, and site-specific integration. The other 20 nucleotides of each ITR, called the D-sequence, although

not part of the hairpin structure, are also necessary for the replication and packaging of the AAV genome (Wang et al. 1997).

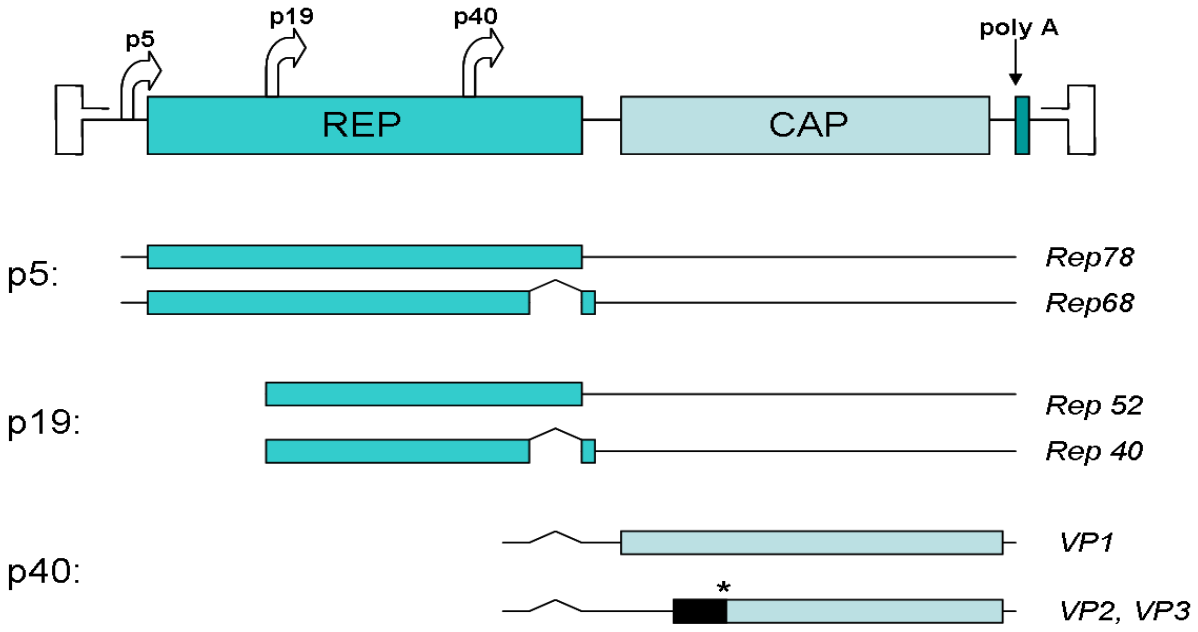



Figure 2-1. The genetic map of AAV. The AAV genome consists of the rep and cap open reading frames (ORF) flanked by the inverted terminal repeats (ITRs). The six mRNA transcripts are shown below the AAV genome map. The larger rep proteins (Rep78/68) are expressed using the P5 promoter. The smaller rep proteins (Rep52/40) use the p19 promoter. The capsid proteins (VP1, VP2, and VP3) are expressed using the P40 promoter. They are produced using alternatively spliced mRNAs. One capsid transcript produces the capsid VP1 subunit. The other transcript produces both VP2 and VP3 subunits: VP2 is translated using a non-conventional ACG start codon, whereas, VP3 is translated using the downstream conventional AUG codon (shown as a star).

Four Rep proteins are encoded by the rep gene: Rep 78, 68, 52, and 40. The larger Rep proteins, Rep 78 and its spliced variant Rep 68, are expressed from the P5 promoter. The smaller Rep proteins, Rep 52 and its spliced variant Rep 40, are expressed from the P19 promoter. All four Rep proteins possess ATPase/helicase and DNA binding activity. In addition, Rep78/68 proteins possess strand/sequence specific endonuclease nickase activity. The Rep proteins are involved in all aspects of AAV biology.

The cap gene encodes three viral proteins, VP1 (87 kDa), VP2 (73 kDa), and VP3 (67 kDa). These particles differ in their N-terminus and share most of the C-terminal sequences. They come together during viral assembly forming a sixty-subunit viral particle at a ratio of 1:1:10 (VP1:VP2:VP3). The viral protein subunits are expressed from the cap gene, using the P40 promoter. Both alternative splicing and alternative start codon usage are involved in translation of the subunits: the VP2 and VP3 proteins are expressed from the same mRNA, a spliced transcript of VP1; VP3 is translated using a conventional start AUG codon, downstream from the non-conventional ACG codon used for VP2.

**AAV Infection**

AAV infection is a multi-step process that includes binding to the cell, endocytosis into cellular vesicles, cytoplasmic trafficking, nuclear entry, viral genome uncoating, and genome conversion for transcription.

**Binding to cellular receptors**

In order to gain entry into target cells, AAV needs to attach to certain cellular receptors and co-receptors. Heparin sulfate proteoglycan has been identified as the primary receptor used by AAV 2 for cell attachment (Summerford et al., 1998). The abundance of heparin sulfate proteoglycan (HSPG) on cellular membranes has accounted for the broad transduction efficiency of AAV 2 observed *in vivo* and in cell culture. A positively charged loop region in the AAV 2 capsid, comprised of residues 585-590 (Opie et al., 2003), has been shown to be necessary for capsid binding to HSPG. For efficient cellular entry, AAV 2 also needs to interact with a co-receptor. It may use one or several of many identified co-receptors including αVβ5 integrin (Summerford et al, 1999), fibroblast growth factor receptor 1 (Qing et al., 1999), hepatocyte growth factor receptor (Kashiwakura et al., 2005), αVβ1 integrin (Asokan et al., 2006), and the 37/67 kDa laminin receptor (Akache et al., 2006) for cellular entry.

Receptors and co-receptors have been identified for other AAV serotypes as well. To date, 13 serotypes have been identified; however the receptor and co-receptor usage has been identified for only a subset of them. Many of the AAV serotypes may use the same receptor, but differences in their co-receptor usage may account for differences seen in their tissue tropisms. The receptors and/or co-receptors have been identified for many of the other AAV serotypes and are summarized in Table 2-2.

Table 2-2. List of AAV serotypes and their identified receptor and co-receptor(s) used for cell entry.

| Serotype | Receptor | Co-receptor | References |
|---|---|---|---|
| AAV 1 | N-linked sialic acid | Laminin receptor | [Chen et al., 2005], [Akache et al., 2006], [Zhijian et al., 2006]. |
| AAV 2 | HSPG | αVβ5 integrin, fibroblast growth factor receptor 1, hepatocyte growth factor receptor, αVβ1 integrin, laminin receptor | In text |
| AAV 3 | HSPG | Laminin receptor | [Akache et al., 2006], [Handa et al., 2000] |
| AAV 4 | O-linked 2, 3 sialic acid | Unknown | [Kaludov et al., 2001] |
| AAV 5 | N-linked sialic acid | Platelet derived growth factor receptor | [Kaludov et al., 2001], [Pasquale et al., 2003] |
| AAV 6 | N-linked sialic acid | Unknown | [Zhijian et al., 2006] |
| AAV 8 | Laminin receptor | Unknown | [Akache et al., 2006] |
| AAV 9 | Laminin receptor | Unknown | [Akache et al., 2006] |

**Internalization and trafficking**

The cascade of cellular events that mediates AAV trafficking is not completely known. After binding, AAV is endocytosed into clathrin coated vesicles, which is regulated by dynamin; dominant negative interference of dynamin abolished AAV transduction, suggesting that AAV enters target cells primarily via clathrin-mediated endocytosis (Duan et al. 1999). Using fluorescently tagged AAV to monitor the internalization and trafficking of AAV in real time,

Seisenberger et al. identified that AAV endocytosis is remarkable fast, with each clathrin vesicle containing a single AAV particle (Seisenberger et al., 2001). AAV has also been observed in non-coated pits, possibly representing a minor form of cellular entry (Bantel-Schaal et al., 2002).

The internalized AAV particles are transported through the cell or are degraded within the endosome-lysosome compartments. The clathrin on the internalized vesicle is removed to allow vesicle fusion, forming a larger compartment called the endosome. The early endosomes go through gradual acidification maturing into late endosomes, which can ultimately fuse with lysosomes, where degradation of the cargo occurs. AAV infection of cells in the presence of the drug bafilomycin A1, an inhibitor of endosome acidification, decreased AAV expression, suggesting that the low pH in endosomes has some role in AAV escape for nuclear entry (Bartlett et al., 2000). The acidification of the vesicles may function to partially disassemble or trigger a conformational change in the AAV 2 capsid while keeping it intact, allowing externalization of VP1 phosphoplipase ($PLA_2$) motif which has been shown to be essential for successful infection (Girod et al., 2002). Mutational analysis has revealed that the $PLA_2$ motif is presented via the 5-fold pore of the AAV capsid during infection (Bleker et al., 2005). Moreover, cathepsin B and L, proteases that reside in the endosomes, can cleave AAV capsids, promoting capsid degradation and possibly also having a role in exposing the $PLA_2$ motif (Akache et al., 2007). Studies using the drugs brefeldin A (inhibitor of early-to-late endosome transition) and MG-132 (proteasome inhibitor) in AAV infection experiments suggest that AAV trafficking to late endosomes is important for infection, and that proteasome-mediated degradation of escaped AAV particles is an obstacle to nuclear entry (Douar e al., 2001).

The process of AAV intracellular trafficking seems to be highly regulated by signaling pathways which are activated upon endocytosis. Inhibiting Notch 1 expression by siRNAs,

prevented AAV internalization in HEK293 cells, whereas, its over-expression increased internalization and perinuclear accumulation of AAV (Ren et al., 2007). Moreover, AAV 2 interaction with integrin αVβ5 during entry resulted in activation of both the GTP-binding Rac1 protein and the PI3K pathway; inhibition of Rac1 using a dominant negative form decreased PI3K activation and AAV trafficking, supporting a role for Rac1 signaling in AAV infection (Sanlioglu, S., Benson, P.K., Yang, J., et al., 2000). The same group reported that inhibition of microtubule and microfilament networks using drugs nocodazole and cytochalasin B, respectively, reduced AAV expression by 90 percent (Sanglioglu, S., Benson, P.K., Yang, J., et al., 2000). The role of microtubules in shuttling AAV has been a bit confounding. In another study, low concentrations of nocadazole treatment had no effect on AAV expression, but at higher levels, it reduced AAV expression at the expense of toxicity to the cells, suggesting that an intact microtubule network may not be required for successful AAV trafficking (Hirosue et al., 2007).

**Nuclear entry**

Once AAV escapes the endosomes, it can enter the nucleus for uncoating and expression. AAV is theoretically small enough (25 nm) to enter through the nuclear pore complex (NPC), yet it remains unclear if it uses the NPC to gain nuclear entry. It has been shown that AAV entry may be partially NPC independent (Hansen et al., 2001; Xiao et al., 2002). In a study by Bartlett et al., Cy3 labeled AAV can be seen at the nuclear periphery by 30 minutes and within the nucleus after 2 hours, suggesting that nuclear entry is a slow step in infection (Bartlett et al., 2000). In another study, coinfection of adenovirus capsids facilitated the nuclear entry of GFP tagged AAV (Lux et al., 2005), supporting previous studies of nuclear entry being rate limiting.

Data from a recent study indicated that mobilization of the AAV particles in the nucleus greatly affects expression; specifically, AAV capsids are sequestered in the nucleolus, and its

movement towards the nucleoplasm, allows uncoating and gene expression to take place (Johnson et al., 2009). This study indicates that sequestration of AAV capsids in the nucleus upon nuclear entry is an obstacle for expression.

Using an *in vivo* footprinting strategy, Wang et al. (2007) demonstrated that rAAV genomes from AAV 2 and AAV 8 capsids formed similar levels of double stranded DNA upon infection, with the onset of expression being quicker using AAV 8. They concluded that the instability of rAAV genomes caused by degradation accounts for the inefficiency of rAAV expression and that kinetics of uncoating may have a role in stabilizing rAAV genomes (Wang et al., 2007). Regardless of which process is rate limiting, nuclear entry or the DNA uncoating, the single-stranded DNA needs to be converted into double stranded DNA in the nucleus, a process which is inefficient (Ferrari et al., 1996) in the absence of helper virus.

**AAV Genome Replication and Expression**

Upon successful nuclear entry and uncoating, AAV expression can take place under two situations, either in the presence or absence of helper virus.

**Presence of helper virus**

In the presence of adenovirus (Ad) or herpes virus, AAV undergoes productive infection characterized by efficient DNA replication, expression, and assembly (Buller et al., 1981). E1a, E1b, E2a, E4, and VA (viral associated) RNA has been identified to be the necessary adenovirus helper genes, and seem to primarily function at the gene regulation level to aid AAV expression (Janik et al., 1981). For example, the AAV P5 and P19 promoters are transactivated in the presence of Ad E1a expression. In its absence, the expression of the AAV genes is negatively regulated in *trans* by Rep (Kyöstiö et al, 1995; Lackner et al., 2002). Adenovirus helper genes can also enhance replication of AAV DNA, which occurs in discrete foci in the nucleus. Cervelli et al. demonstrated that AAV interacts with the cellular MRN (Mre11-Rad50-NBS1) complex at

these foci which inhibits AAV replication (Cervelli et al., 2008). This inhibition can be relieved by adenoviral E1b55k/E4orf6 expression (Schwartz et al., 2007). The other adenoviral genes (E2a and VA RNA) may function to increase or stabilize AAV mRNA levels for translation.

Herpes virus genes primarily regulate AAV replication. For example, herpesvirus helicase complex proteins (UL5/8/52) and single stranded DNA binding protein (UL29) have been shown to enhance AAV replication (Mishra et al., 1990; Weindler et al., 1991; and Ward et al., 2001). In addition, HSV1 ICP0 protein, although not necessary as a helper function, enhances the expression of the AAV Rep gene by regulating the Rep promoter (Geoffrey et al., 2004). Despite the differences in helper function between adenovirus and herpesvirus, both seem to provide a permissive cellular environment for a productive AAV infection.

AAV replication is an essential first step for productive infection, which occurs in the presence of helper virus genes. The ITR serves to provide a free 3' OH end onto which the DNA polymerase synthesizes a second strand to form a linear double stranded DNA, called a replicating form monomer (RfM). A second round of self-priming synthesis is initiated using the RfM, forming a linear double stranded dimer DNA, now called a replicating form dimer (RfD). These double stranded intermediates are then processed via a strand displacement mechanism, resulting in single stranded DNA forms, which are used for packaging, and double stranded DNA forms, which are used for transcription and the next round of replication.

Critical for resolution of the double strand intermediates (RfM and RfD) are *cis* sequences in the ITR which include the Rep binding elements (RBE, RBE') and a terminal resolution site (TRS) (Snyder et al., 1993) (Figure 2-2). The larger rep proteins, Rep68/78, form a hexamer on the ITR hairpin (Li et al., 2003), and through interaction with the RBE elements is functionally

situated to initiate efficient nicking of the TRS between the two thymidine bases, a process called terminal resolution (Brister et al., 2000). This allows for re-initiation of replication allowing

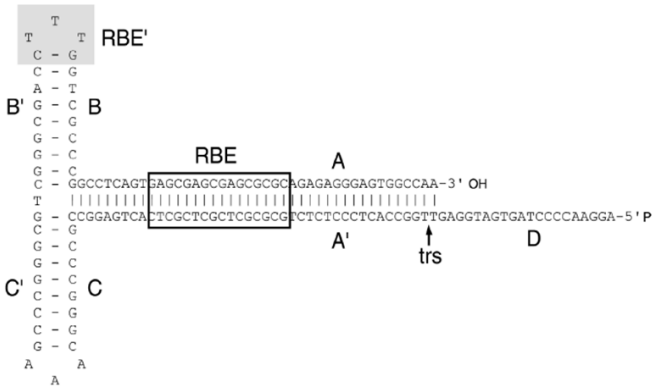



Figure 2-2. Elements in the AAV ITR required for replication. The Rep binding element (RBE) sequence is GAGCGAGCGAGCGCGC. The RBE' sequence is CTTG. The terminal resolution site (TRS) sequence is GTTGG; nicking occurs between the thymine bases.

strand displacement. The Rep78/68 proteins necessary for AAV replication possess ITR binding activity, helicase activity, ATPase activity, and site/strand-specific endonuclease activity at the terminal resolution site (TRS) (Chiorini et al., 1994). The smaller Rep proteins, Rep 52 and Rep 40, are important for the accumulation of the single-stranded AAV genome generated during strand displacement to be used for packaging (Chejanovsky et al., 1989). Using nuclear extracts from Ad-infected cells and an *in vitro* AAV replication assay, Nash et al. identified replicating factor c, proliferating cell nuclear antigen, DNA polymerase delta, and the minichromosome complex proteins as necessary cellular factors for AAV replication (Nash et al., 2007; Nash et al., 2008). In addition, AAV replication can be enhanced in the presence of high mobility group 1 which enhances Rep's DNA binding and nicking activity (Costello et al., 1997)

During productive infection replicated AAV DNA is packaged into newly assembled AAV capsids. The precise mechanism of AAV assembly is not known. However, some basic discoveries have provided insight into the AAV assembly process. Several groups have

described the interaction of the Rep78/68 proteins with the capsid suggesting that they are involved in a complex that brings the DNA near empty capsids for packaging (Prasad et al., 1995; Wistuba et al., 1995). The residues important for this interaction are present in all four Rep proteins (Dubielzig et al., 1999). Further studies have identified that Rep 52 and 40 can actively translocate the single stranded AAV genome into preformed empty AAV capsids in 3' to 5' fashion (King et al., 2001). Mutational alteration of the AAV capsid five-fold pore affected both packaging and Rep-DNA interaction, providing evidence for a role for the five-fold pore as the site of DNA insertion into preformed capsids (Bleker et al., 2006). Grieger et al. identified two basic capsid regions, BR1 (residues 166-172) and BR2 (residues 307-312), as being essential for virus assembly (Grieger et al., 2006). These regions are well conserved among the AAV serotypes, supporting their significance in the AAV assembly mechanism.

**Absence of helper virus**

The process of AAV expression is remarkably different in the absence of helper virus. First of all, the expression of Rep is negatively autoregulated. Both Rep expression and the presence of the P5 RBE are necessary to transactivate the P40 promoter which drives expression of cap, thereby, regulating capsid expression as well. The presence of adenovirus gene expression relieves this repression, allowing both rep and cap genes to be expressed. Moreover, a cellular transcriptional protein, PC4, has been identified as a rep-binding partner. Rep can bind both the transcriptionally active (non-phosphorylated form) and inactive forms (phosphorylated form) of PC4; and in the absence of helper virus genes, PC4 represses transcription from all four AAV promoters (Weger et al., 1999).

Other cellular factors have also been identified to have a role in AAV transduction and expression. FKBP52, a nuclear protein, was found to bind the D sequence in the AAV ITR (Ashktorab et al., 1989; Qing et al., 2001). The phophorylation state of this protein modulates

the binding to the AAV DNA, which can prevent AAV double strand DNA (dsDNA) formation (Qing et al., 2001). Inhibition of epidermal growth factor receptor protein tyrosine kinase (EGFR-PTK) and over expression of T-cell protein tyrosine phosphatase (TCPTP) augmented AAV expression by reducing the binding of FKBP52 (Mah et al., 1998; Qing et al., 2003). Moreover, protein phosphatase 5 (PP5) can phosphorylate FKBP52 at serine and threonine residues, also regulating the binding state of FKBP52 (Zhao et al., 2006). Taken together, these results indicate that in the absence of helper virus genes, the AAV genome is acted on by cellular proteins that inhibit AAV second strand synthesis and expression. Several groups have provided evidence that second-strand synthesis is the primary means for dsDNA formation of rAAV genomes for expression (Zhong et al., 2008; Zhou et al., 2008).

A recent study showed that EGFR-PTK can also phosphorylate the AAV 2 capsids, marking them for ubiquitin-mediated degradation via the proteasome pathway. This effectively reduces the number of AAV particles that make it into the nucleus for uncoating and expression (Zhong et al., 2007). Moreover, the same group reasoned that mutating the tyrosine residues on the AAV capsid surface can reduce proteasomal mediated degradation of AAV capsids, thus augmenting AAV nuclear entry for expression. Using site-directed mutagenesis, they were able to create novel AAV tyrosine-mutant AAV vectors that showed greater transduction properties both *in vitro* (in HeLa cells) and *in vivo* (in murine hepatocytes) at a log lower dose, compared to wild type AAV capsid vectors (Zhong, L., Li, B., et al., 2008).

## Recombinant AAV Vectors for Gene Therapy

Viruses are pathogens that can cause disease in humans. They have an innate ability to infect human cells and tissues. The use of these viruses as vectors to deliver therapeutic genes has revolutionized the gene therapy field. Ideal vectors retain the virus's ability to infect, but

without the pathogenicity. This is done by removing the viral genes that cause disease, and retaining only those elements required for expression and packaging.

**General Description of AAV Vectors**

AAV was first used as a vector in 1984 to transduce cultured cells (Hermonat et al., 1984). Since then, it has achieved considerable popularity as a vector because it has not been associated with disease and is a naturally defective virus. The most common AAV virus used for vector development is AAV 2. Recombinant AAV (rAAV) vectors are created by removing the rep and cap genes. The ITRs are maintained since they are required for packaging. rAAV vectors are promising, because they display long term expression in human tissues, where cells are largely post-mitotic. In dividing cells, rAAV is rapidly diluted and doesn't readily integrate into the human genome, like retroviruses, or wild type AAV (Rep positive). The limitation of rAAV vectors development is the small packaging capacity (4.7 kb) for transgene insertion. AAV-based vectors display persistent expression in post-mitotic tissues, have no or low immunogenicity, the different AAV serotypes allow for tissue specific targeting, and AAV can infect both dividing and non-dividing cells.

**Production and Purification of AAV Vectors**

The popularity of rAAV vectors have prompted the demand for increased characterization of production and purification approaches. The broad use of rAAV vectors for gene transfer has been hampered by production and purification obstacles that affect its purity and titer. These include, (i) concerns over the dependence on adenovirus for production, and the contamination of final rAAV stocks with infectious adenovirus, (ii) reliance on density gradients for purification which makes large scale purification difficult and (iii) lack of standardized titering approaches to determine potency of purified rAAV vectors.

**Production of AAV vectors**

Approaches for vector production that have been studied include, (i) using stable producer cells expressing rep and cap genes which requires only the input of the rAAV plasmid and Ad infection for production (Tamayose et al., 1996), and (ii) a transfection approach that removes the requirement for adenovirus infection (Salvetti et al. 1998; Xiao et al., 1998). Some interesting observations have improved rAAV production and yield. First of all, the discovery of the Ad genes (E2a, E4, VA I RNA, E1a, E1b) required for AAV replication, allowed the removal of the adenovirus infection from the production process. These essential genes could be incorporated into a plasmid and when used in rAAV production, it improved rAAV yield and the final rAAV stocks were free of infectious adenovirus (Xiao et al., 1998). Moreover, transfecting a 1:1:1 molar ratio of the three plasmids (rAAV plasmid containing the ITRs, Ad-mini gene plasmid, and AAV helper plasmid containing the rep and cap genes) provided an optimal yield of encapsidated rAAV vectors (Xiao et al., 1998). The second observation was that downregulation of Rep78/68 expression from the AAV helper plasmid increased rAAV yield 10 fold (Li et al., 1997), by improving rAAV DNA replication and expression of the cap gene. Last of all, the development of a two-plasmid system consisting of the rAAV vector and the Ad-AAV plasmid, simplified the transfection process and increased efficiency of rAAV production. For example, Grimm et al. described a two-plasmid system where they constructed different versions of the Ad-AAV plasmid, pDG, allowing for cross-packaging of AAV 2 genome into capsids of different serotypes. To allow identification of the serotype-specific capsid, they incorporated different fluorescent markers into the plasmids for visual identification (Grimm et al., 2003).

Other production methods are being developed that aim to produce higher rAAV virus yields required for clinical applications. These include the use of producer cells that only need helper virus infection (i.e herpesvirus) and the production of rAAV in baculovirus (insect) cells

(Sollerbrant et al., 2001; Booth et al., 2004; Merten et al., 2005). These methods reduce or eliminate the requirement for transfection, and can be scaled-up for clinical production levels, thus improving the production yield of packaged rAAV vectors.

**Purification of AAV vectors**

Currently, on laboratory scale, rAAV is packaged in HEK 293 cells using the triple or double transfection approaches described above. The cells are typically harvested at 48-72 hours post-transfection, allowing for maximum rAAV virus production. Recombinant virus is retained within the cells, therefore, physical methods are needed to lyse the cells and remove the produced rAAV virus. This is done by lysing the cells in phosphate buffered saline using three cycles of freeze (15 minute on dry ice) and thaws (15 minutes at 37°C). The freeze-thaw cycles efficiently release most of the packaged recombinant virus from the cells into the supernatant. The cellular debris is removed by centrifugation, and the rAAV containing supernatant is treated for further purification. Specifically, benzonase is used to digest cellular and non-encapsidated DNA and/or detergents are used to prevent cellular protein aggregation to viral particles. Next, the recombinant virus in the treated supernatant is purified using cesium chloride gradients, iodixanol centrifugation, and affinity column, or ion-exchange column chromatography.

The traditional method of rAAV purification involved using multiple rounds of cesium chloride centrifugation. Problems with cesium chloride purification include its toxicity, its inability to remove excess cellular contaminants, difficulty in scaling up the purification process for human application, and the high ratio of purified non-functional to functional rAAV virus in final stocks. The main advantage of cesium chloride purification is it allows for separation of empty capsids from rAAV packaged capsids.

More recent purification approaches have been developed such as iodixanol gradient centrifugation, ion-exchange columns, and affinity columns. Iodixanol based gradients have been

developed because iodixanol is non-toxic and the gradient allows the separation of cellular contaminants from rAAV particles. For purifying rAAV 2, the virus is pulled from the iodixanol gradient and loaded onto a heparin column. By combining gradient and column purification, over 50% of the viruses from the supernatant can be recovered and are over 99% pure. The heparin column can also been used to purify rAAV 3 vectors. AAV 4 and AAV 5 can be purified used mucin columns, based on the affinity of these serotypes for sialic acid.

Ion-exchange chromatography approaches for AAV purification has been the most promising with regards to purity and scalability. All the AAV serotypes can potentially be purified using ion-exchange, however, the salt and pH conditions need to be optimized for each serotype, since each one possesses unique capsid surface charges. Under the right conditions the AAV particles can efficiently bind to the column matrix and the weakly interacting contaminants can be washed away. The bound virus can then be eluted off these columns using a salt gradient. The fractions containing most of the virus are pooled and concentrated. Ion-exchange chromatography has been used to purify AAV serotype 2, 4, 5 and 8. Compared to conventional cesium chloride purification and affinity columns, ion-exchange columns allows maximum recovery of recombinant virus from supernatants, results in highly pure rAAV stocks, and it can be easily scaled up for human application.

A small aliquot of the purified, concentrated rAAV virus is used to quantitate the amount of virus recovered. The physical amount of virus (viral particles per milliliter) can be determined, for AAV 2 only, using a commercially available A20 antibody-based enzyme linked immunosorbent assay (ELISA). The number of rAAV genome containing particles per milliliter is determined using PCR or dot blot hybridization using the encapsidated rAAV DNA. The last titering method involves determining the number of transducing viral particles per milliliter – the

proportion of viral particles that can infect and express in cells. Among these titering approaches, the last one is the ideal because it determines the amount of functional virus per milliliter of purified rAAV stock.

## New AAV Vectors

### Self-complementary AAV vectors

Self-complementary AAV (scAAV) vectors have been developed to bypass the limiting aspect of second-strand synthesis (McCarty et al., 2001), previously described. Generation of rAAV vectors depends on replication and resolution of the replicated forms (RfM and RfD). Failure to resolve these replicated forms, leads to accumulation of the dimer form (RfD) of the AAV DNA. Therefore, by deleting the terminal resolution site of one ITR, Rep cannot fully resolve the RfD, resulting in greater packaging efficiency of dimeric genomes (McCarty et al., 2003), called self-complementary vectors.

Upon transduction and uncoating, scAAV vectors can fold upon themselves, immediately forming transcriptionally competent double-stranded DNA. These vectors increase the onset of expression by bypassing the need for dsDNA formation. One consequence is that the vector capacity is reduced, but up to 3.3 kbp of DNA can be encapsidated (Wu et al., 2007), limiting the capacity for insertion of regulatory elements in the vector. Rapid expression has been observed in many tissues *in vivo* and in cell cultures. Production and purification approaches for scAAV vectors are the same as they are for rAAV vectors. However, the yield of scAAV production can be greatly affected by the expression of Rep: excess production of Rep78/68 reduces the percentage of total particles containing scAAV genomes, presumably due to packaging of single stranded DNA (ssDNA) genomes generated by spurious resolution of dimeric genomes into monomeric forms.

Self-complementary AAV vectors (scAAV) are not recognized within the transduced cells the same way as single stranded AAV vectors (ssAAV) by host proteins. But, similar to ssAAV vectors, scAAV vectors can also form episomes. Using different DNA-repair deficient cells lines, Choi et al. demonstrated that RecQ helicase family members (BLM and WRN), Mre11, NBS1, and ATM are required for scAAV circularization. *In vivo*, DNAPKcs and ATM were required for scAAV circularization; however, NBS1 was dispensable (Choi et al., 2006). Moreover, intermolecular recombination between two scAAV vectors occurs at both ends with different efficiencies (Choi et al., 2005). It will be important to further characterize the molecular organization of scAAV vectors and their persistence in transduced tissue, to gain a better understanding of how scAAV vectors can be developed further for gene therapy applications.

Many clinically relevant tissues remain refractory to transduction and expression by AAV 2, even when using scAAV vectors. Some of these tissues have been efficiently transduced using the other available AAV serotypes. Skeletal slow and fast muscle fibers are transduced efficiently with AAV 2 vectors packaged into AAV 7 (AAV 2/7) and AAV 8 capsids (AAV 2/8) (Louboutin et al., 2005). Moreover, muscle tissues can also be efficiently transduced with AAV serotype 1, 6, 7, 8, and 9 (Gao et al., 2002; Chao et al., 2000; Blankinship et al., 2004; Wang et al., 2005; and Pacak et al., 2006). Greater gene expression has been observed in neurons of murine brain using AAV 1, which can also transduce glial and ependymal cells, unlike AAV 2 (Wang et al., 2003). AAV 5 has preference for retina (Rabinowitz et al., 2002) and joint cartilage (Apparailly et al., 2005). Compared to AAV serotype 2, AAV vectors based on serotype 8 and 9, display superior transduction in liver (Davidoff et al., 2004; Gao et al., 2002; Nakai et al., 2005;

Inagaki et al., 2006). Hematopoietic cells, which are refractory to AAV 2 infection, can be efficiently transduced with vectors based on AAV 3 (Handa et al., 2000)

The use of the different AAV serotypes in a pseudotyping approach (the genome of one serotype packaged into a different serotype capsid) to increase the transduction of AAV in different tissues, coupled with use of self-complementary AAV vectors (scAAV), has greatly improved expression in many human tissues.

**Trans-splicing vectors**

An additional novel AAV vector system has been developed, that takes advantage of the ability of AAV to form multimers, to increase the genome capacity of AAV (Yan et al., 2000). This system relies on AAV's ability to form head to-tail concatemers via ITR recombination. In this approach, the transgene cassette is split between two rAAV vectors containing adequately placed splice donor and acceptor sites. Transcription from recombined rAAV genomes, followed by the correct splicing of the mRNA transcript, will result in the expression of a functional gene product. This application is useful for delivering genes up to 9 kb in size. Trans-splicing has been successfully used in the retina (Reich et al., 2003), lung (Liu et al., 2005), and muscle (Ghosh et al., 2006; Ghosh et al., 2008). In terms of efficiency of expression, only a small percentage of input vectors recombine, therefore, trans-splicing vectors are less efficient than rAAV vectors. These vectors have been very useful for expressing CFTR transgene (7.2kb), Dystrophin (6kb), and Factor IX cassettes (7kb). The efficiency of trans splicing vector expression can be increased by using vectors containing ITRs of different serotypes at each end, respectively. These vectors are less likely to circularize and more likely to form linear concatemers, promoting efficient vector dimerization for expression (Yan et al., 2005).

**Fate of Recombinant AAV Vectors**

An early study of rAAV vectors *in vivo* had concluded that rAAV exist primarily as episomes, which could be rescued upon Ad super-infection (Afione et al., 1996). It is believed that rAAV random integration occurs at a remarkably low frequency (10^-7). One study failed to find integrated forms of AAV in transduced muscle tissues (Schnepp et al., 2003). In addition, skeletal muscle transduction with rAAV vectors resulted in stable expression detected at 2 weeks post-injection, which came from rAAV genomes that were in the form of high molecular weight concatemers (Vincent-Lacaze et al., 1999). Further studies to characterize the molecular forms of the rAAV in transduced muscle, identified that they were arranged into large (>12 kbp in size) circular head- to-tail multimers, which were initially monomers that slowly converted to larger forms via ITR recombination over time (Duan et al., 1998). This circularization of rAAV genomes was shown to be regulated by DNAPKcs *in vivo*: in the absence of DNAPKcs more linear rAAV genomes were present in muscle tissue compared to controls (Duan et al., 2003).

Concatemer formation of rAAV genomes may occur via either DNA replication or through intermolecular recombination. The analysis of the episomal structures in the above studies provided strong evidence for the latter mechanism. Often during inter or intramolecular recombination, deletions occur in the ITRs, with the head-to-tail form predominating, although head-to-head and tail-to-tail forms of rAAV recombination have been observed (Yang et al., 1999). In another study, analyses of injected double-stranded circular and linear rAAV genomes containing or missing ITRs from mouse livers demonstrated that circular genomes are not processed into larger concatemers (Nakai et al., 2003).

The fate of rAAV genomes and their concatemers as pre-integration intermediates has not been fully characterized. Using an AAV 2 based shuttle vector to isolate rAAV-cellular sequences in vivo, Inagaki et al. identified that rAAV vectors were able to integrate into the

human genome in post-mitotic tissues; and using partial hepatectomy, they estimated that around 0.2 rAAV genomes had integrated per diploid genome in livers of neonatal mice (Inagakai et al., 2008). Analysis of integration in post-mitotic mice tissues such as skeletal muscle and heart after rAAV gene transfer revealed that palindromic regions in the human genomes are preferential targets for rAAV integration (Inagaki,K., Lewis, S.M., et al., 2007). Together these studies suggest that rAAV genomes, like wild type AAV, persist largely as episomes in post-mitotic human tissues. rAAV and wild type AAV genomes are still capable of random integration *in vivo*, of which the frequency is not known but presumed to be very low

# CHAPTER 3
# RESEARCH OBJECTIVES

## Background

### AAV Site-Specific Integration

The observation that AAV has the ability to establish latency by integrating its genome into the host DNA is of central importance for gene therapy treatments requiring stable, long term, and safe expression. A better understanding of the mechanism and efficiency of site-specific integration, and host factors requirements is needed to create novel, AAV site-specifically integrating recombinant vectors for gene therapy applications.

Early experiments in characterizing AAV latency were performed by Berns et al. who reported that Detroit 6 cells infected with 250 infectious units of wild type AAV 2 virus maintained AAV sequences after 50 passages. Moreover, they demonstrated that roughly 30% of clones generated from these infected Detroit 6 cells were able to rescue AAV sequence upon subsequent adenovirus super-infection (Berns et al., 1975). The first evidence for AAV establishing latency by integration was provided by Cheung et al. who analyzed the genomic DNA from a latently infected Detroit 6 cell line and found that wild type AAV was oriented into a head to tail configuration and associated with high-molecular weight DNA via junction with the ITRs (Cheung et al., 1980). Further characterization of AAV provirus sequences identified that junctions frequently occurred at or within the terminal repeats with the flanking cellular sequence being amplified, suggesting that some DNA replication takes before or during the integration process (McLaughlin et al., 1988).

Several groups later presented direct evidence that AAV establishes latency by integrating into a common site: Chromosome 19q13.4-qter (Kotin et al., 1990; Samulski et al., 1991), a site termed AAVS1. Using an EBV shuttle vector, which can be stably maintained as episomes in

eukaryotic cells and recovered in bacteria, containing cloned different sized AAVS1 fragments (8.2 kb, 4.4 kb, 3.5 kb, 1.6 kb, and 0.510 kb), Giraud et al. demonstrated that a 510 nucleotide sequence at the 5' end of the AAVS1 region is required for efficient recombination between the EBV containing AAVS1 vector and the AAV genome, with roughly 0.8 % of rescued EBV plasmids hybridizing to AAV sequences (Giraud et al., 1994). The recombination signal was narrowed down to a 33 nucleotide AAVS1 sequence consisting of a Rep binding element (RBE: GCTCGCTCGCTCGCTG), and a terminal resolution site (TRS: GGTTGG) separated by an eight nucleotide spacer sequence. Mutating the AAVS1 RBE or TRS sequences completely abolished integration (Linden et al., 1996). The spacer sequence can be altered to some extent but removing or mutating a central CTC sequence reduced integration (Meneses et al., 2000). *In vitro* data by Weitzman et al. demonstrated that Rep68/78 can bind both AAV and AAVS1 RBE sequences in an *in vitro* system, suggesting a mechanism of integration by which Rep can bring these two sites together (Weitzman et al., 1994), a pre-requisite for replication-based integration dependent on Rep68/78 (Urcelay et al., 1995). A PCR-hybridization based approach was developed and used by Surosky et al. to demonstrate that only the larger Rep proteins (Rep78 and 68) can target site-specific integration, which can be enhanced by the presence of the ITRs, although not necessary for integration (Surosky et al., 1997).

Interestingly, Rep is flexible in its binding specificity. Wonderling et al. identified eighteen new Rep binding sites which are generally located in or near genes, and used electrophoretic mobility shift assay to show that these new sites do bind Rep in *vitro* (Wonderling et al., 1997). Interestingly, an estimated 200,000 potential binding sites may exist in the human genome based on the identified minimum GAGYGAGC Rep binding sequence (Young et al., 2000). These data suggest that there may be alternative integration sites for AAV that may also be Rep-mediated.

But, the AAVS1 locus is the only site that contains both a RBE and TRS, underlining the importance of both binding and nicking by Rep in the integration process.

**AAVS1 and Its Characteristics**

The AAVS1 region is very rich in overall guanine and cytosine content (65% GC), with the first 900 bases fulfilling the criteria of a CpG island (82% GC) (Kotin et al., 1992). AAVS1 itself is a 4.7 kb part of a gene called protein phosphatase 1 regulatory subunit 12C or MBS85 (myosin binding subunit 85) thought to be involved in regulation of actin-myosin assembly (Tan et al., 2001). A DNAse hypersensitive site (DHS-S1) has been identified in the AAVS1 CpG island, suggesting that AAVS1 is open to transcription and Rep protein binding (Lamartina et al., 2000). An insulator has also been identified near the DHS-S1 and shown to prevent spread of expression by the enhancer activities of the DHS-S1 (Ogata et al., 2003).

Several muscle specific genes including TNNT1 (slow skeletal muscle troponin T) and TNNI3 (cardiac troponin I) have been identified near AAVS1 (Dutheil et al., 2000). Interestingly, AAV specific integration promotes partial duplication of the AAVS1 locus, preserving a functional copy of the MBS85 gene; integration didn't affect expression of the nearby muscle specific genes or that of MBS85 (Henckaerts et al. 2009).

The AAVS1 integration site has been identified in non-human primates (Amiss et al., 2003) and a mouse AAVS1 ortholog has been characterized (Dutheil et al., 2004). The recombination signals (RBE, spacer, and TRS sequences) in non-human primates have 98% homology to the human site, with the simian RBE sequence being extended by an additional GAGC sequence. Similar to the human site, the mouse ortholog recombination signals can be bound and nicked by Rep, whilst sharing only 74% sequence homology.

**Mechanism of Rep-mediated integration**

Analysis of cloned AAV-AAVS1 junctions suggests that non-homologous recombination is taking place, since only small homologies are seen at the junctions. In addition, deletions in the ITR suggest that significant DNA processing occurs during integration. Several AAV site-specific integration models have been proposed. They all have some common events they share: 1) the AAV genome is brought in close proximity to AAVS1 via Rep binding to the RBE in both DNA sequences, 2) this is followed by nicking of the TRS site in AAVS1 forming a free 3' OH end and covalent 5' Rep-DNA complex, 3) unwinding of the host duplex DNA, and the initiation of DNA replication that starts on the unwound AAVS1 DNA, and 4) through several rounds of template switching the polymerase incorporates the AAV sequence into AAVS1. Concurrent with this process is the gap filling and DNA processing by other cellular factors. It has been proposed that since the integrated AAV is often arranged in a tandem arrangement, that the AAV genome is circularized before integration.

**Elements required for integration**

As stated above, the RBE and TRS sites are essential. A 138 bp P5 sequence, termed the integration efficiency element (P5IEEE), has been shown to enhance integration by 10-100 fold (Philpott et al. 2002). Feng et al. tested a series of constructs containing various lengths of P5IEE and found that the 16 bp RBE was necessary and sufficient for integration and interestingly, noticed that the RBE from the ITR was more effective at integration (Feng et al. 2006). These observations are a bit confounding, as another group identified that the entire P5 sequence is necessary for integration and the ITR RBE did not mediate better integration efficiency (Murphy et al., 2007).

Dyall et al. used an in vitro system to generate AAV-AAVS1 junctions and found that the junction formation is highly dependent on ATP suggesting that the helicase activity of Rep is

important (Dyall et al. 1999). Interestingly, HMG1, a DNA binding protein, has been shown to bind Rep and enhance its endonuclease activity (Costello et al., 1997). Whether HMG1 also enhances Rep mediated TRS nicking on the AAVS1 hasn't been demonstrated. Isolation of AAV-AAVS1 junction from clinically relevant tissues has been difficult. Interestingly, one study found evidence of AAV specific integration in human testis tissue, the significance of which is unclear at the moment (Mehrle et al., 2004). Numerous others studies have failed to find AAV specific integration (Schnepp et al., 2005; Schnepp et al., 2009), suggesting that the frequency of specific integration *in vivo* is very low.

**Efficiency of integration**

It has been very difficult to estimate efficiency of site-specific integration because it is highly dependent on the virus multiplicity of infection (MOI) being used for infection, type of cells being infected, and the methods used for integration analysis. A quantitative PCR of AAV-AAVS1 junction has been used to compare integration efficiency in total infected cells. Based on this approach around 10-20% of AAV infected cells have AAV integrated in the absence of selection and an estimated 0.1% of infecting viral genomes is capable of integration (Huser et al., 2002). Another approach has been to clone out individual cells after infection for analysis by genomic southern hybridization. Based on this approach, integration efficiencies in the range of 20 to as great as 70%, has been observed. This broad range is likely due to difference in cell type used, MOI infection, and cloning efficiency. Consistent among many studies, maximal site-specific integration is postulated to takes place somewhere between 24 to 48 hours post-infection. Interestingly, the stability and expression of site-specifically integrated transgenes seems to be dependent on whether Rep is provided in *cis* or *trans*. In one study site-specifically integrated cell lines generated using Rep in *cis* showed declining transgene expression and loss of transgene DNA; in contrast, cell lines generated when Rep was provided in *trans*, maintained

the transgene DNA and the expression throughout the experimental 18 weeks (Philpott et al., 2004).

AAV is a unique mobile genetic element. It possesses site-specific integration ability and the integrated genome can be rescued and replicated upon adenovirus super-infection. Whether Rep-mediated integration of AAV is ready for clinical applications remains to be evaluated. Two approaches for gene therapy can be envisioned: (i) in the absence of Rep how does AAV persist apart from being an episome. Is it able to integrate, where, and at what frequency? and (ii) in the presence of Rep, how does the distribution shift to site-specificity?, and are there other integration sites that Rep can target? It has been shown that rAAV vectors, without Rep, do not integrate site-specifically (Kearns et al., 1996; Ponnazhagan et al., 1997). Unfortunately, a lot remains to be discovered with regards to the mechanisms of site-specific integration and the host-requirement that mediate integration.

## Research Objectives

The two objectives of my thesis will serve to provide some fundamental basis for understanding more about the biology of AAV latency.

### Determine the Distribution of Rep-mediated AAV Integration

The first research objective was to design and test an experimental approach that can be used to determine the specificity of AAV site-specific integration. Specifically, the aim looked at the distribution of AAV integration in the presence of the AAV replication protein (Rep), to determine if Rep can minimize random integration events or simply promote more frequent site-specific integration. An AAV shuttle vector was developed which would allow rescue of integrated sequences as plasmids in bacteria. In addition, a genome walking kit was used that relies on PCR amplification to isolate integrants with the flanking cellular sequences. It was

anticipated that these approaches would help in answering the question of whether Rep can minimize random integration events.

**Determine the Role of Cellular DNA Repair Proteins on Site-Specific Integration**

This research objective describes the studies that were aimed at determining if cellular repair proteins (DNAPKcs, Ligase I, and Ligase IV) have an effect on AAV site-specific integration. Specifically, an emphasis was placed on studying how DNAPKcs affect site-specific integration, since a previous report has indicated that it inhibits recombinant AAV integration (Song et al., 2004). The frequency of site-specific integration was assessed for both single-stranded and self-complementary AAV vectors in the absence or presence of DNAPkcs. It is anticipated that an expansion on my initial investigations should shed more light into the biology of AAV site-specific integration and can provide a framework from which to assess the safety of integration and to develop novel site-specifically integrating AAV vectors.

# CHAPTER 4
# DETERMINE THE DISTRIBUTION OF REP-MEDIATED INTEGRATION

## Introduction

Major hurdles to the use of viral vectors for gene transfer have been the host response to the viral vector and the risk of insertional mutagenesis associated with random integration of the delivered viral vector. Several studies have suggested that integration of recombinant AAV vectors is an important issue to address, since it can have consequences for safe gene expression. The first direct evidence for rAAV integration was presented by Nakai et al., whom infected mouse liver with a rAAV shuttle vector and analyzed integration in genomic DNA from mouse livers 5 months post-injection using a bacterial trapping approach. Analysis of 18 junctions revealed deletions and amplification of both the AAV vector and cellular sequences, with integration occurring in several genes (Nakai et al., 1999).

Miller et al. performed a large scale *in vitro* analysis of rAAV integration sites in normal human fibroblasts in the absence of selection (Miller et al., 2005). They also used an AAV shuttle vector to rescue proviral sequences containing flanking cellular DNA in bacteria as circular plasmids. Around 977 unique AAV-cellular junctions were analyzed with rAAV found integrated into all 23 chromosomes. Similar to the *in vivo* study by Nakai et al., deletions and amplifications were observed. Interestingly, when the distribution of rAAV integration was compared to a randomized control set, it became obvious that rAAV vectors have preferences for CpG islands and ribosomal repeats, with a modest preference for transcriptional units. Although none of the integrations sites included known oncogenes, the conclusion was that that rAAV integration is taking place at sites prone to DNA breaks.

It seems that rAAV vectors have a broad distribution of integration in human cells. Unlike other integrating vectors, rAAV vector integration seems relatively uncommon, probably no

more efficient than integration of transfected naked DNA. It is important to note that rAAV vectors do not encode viral proteins capable of causing host DNA breaks; therefore, the consensus is that rAAV integration is taking place at site of existing chromosomal breaks. Although this raises concerns over rAAV integration that may cause tumorigenesis, rAAV integration has not been associated with cancer. The most likely concerns regarding rAAV integration would stem from the chromosomal deletions and amplifications associated with integration and its effect on cellular function.

The mechanism of rAAV integration is not known, but it is likely that cellular nuclease processing of DNA damage sites expose host DNA sequences that can form microhomologies with processed rAAV vectors, resulting in integration. The imprecise mechanism and chromosomal changes suggest the involvement of non-homologous end joining (NHEJ) proteins.

The correction of many diseases require long term transgene expression, making integrating vectors ideal candidates. However, the risk of insertional mutagenesis and oncogene activation are important considerations. Therefore, the identification of rAAV integration in or near genes suggests that rAAV vectors may present safety issues for gene transfer. However, an alternative does exist: the use of Rep containing AAV vectors which can integrate site-specifically. In the presence of the AAV Rep protein the AAV genome is targeted for integration into chromosome 19 site, AAVS1. Rep possesses an endonuclease function and has been shown to nick the AAVS1 chromosomal DNA at the TRS. Unlike retroviruses that can nick host DNA throughout, Rep-mediated nicking may just occur at AAVS1, the only site with a nicking sequence (TRS) in close proximity to a rep biding site (RBE). The distribution of rAAV random integrations in the absence of Rep suggests that sites of integration occur at spontaneous chromosomal breaks. Interestingly, the distribution of wild type AAV hasn’t been analyzed in

detail. Of interest would be to see if Rep, which is used to target wild type AAV site-specific integration, can minimize random integration or simply promote more frequent specific integration. Therefore, my first aim attempted to identify the distribution of integration of an AAV Rep containing shuttle vector or wild type AAV, using several different molecular techniques.

## Experimental Design and Methods

### Construction of an AAV Vector, P5-Rep Shuttle.

An AAV shuttle vector, P5RepShuttle (4.7 kbp), was constructed to allow direct isolation of AAV integrants. Using standard cloning procedures, a 2.4 kb insert containing the left ITR, P5 promoter, and Rep was cut out from pAV2 (wild type AAV 2 vector). The insert was ligated to a 4.1 kb backbone isolated from an AAV-EF1a-GFP.AOSP vector (a kind gift from Dr. Nakai). A 200 bp PCR product (from P5UF11) containing a poly A sequence was inserted at an NdeI site near the Rep gene.

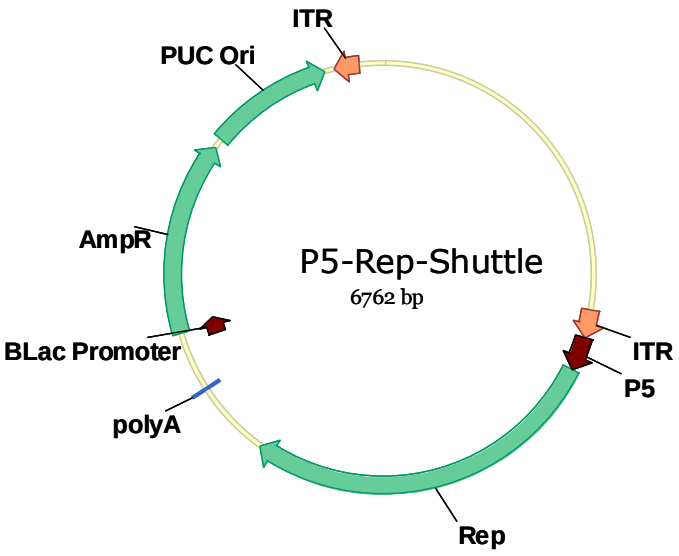



Figure 4-1.Map of P5RepShuttle Vector. The shuttle vector consists of the P5 promoter driving Rep78 expression. Included are the Amp and pUC Ori sequences for bacterial trapping experiments. Flanking all these elements are the inverted terminal repeats.

The key features of this vector (Figure 4-1) are: (i) A P5 promoter containing a 138 bp integration efficient element (IEE) that was demonstrated to increase AAVS1 integration 100 fold (Philpott et al., 2002), (ii) the Rep gene, which is essential for targeted integration, (iii) an Ampicillin resistance gene for selection in bacteria, and (iv) a bacterial plasmid origin of replication for amplification in bacteria, all enclosed within two inverted terminal repeats (ITRs).

**Western Blot for Rep Expression from P5RepShuttle Vector**

Whole cell proteins were isolated using Cell Lytic Reagent (Sigma), according to manufacturer's instructions. The proteins were loaded onto a 7.5% Tris-HCL precast SDS-Page gel (Bio-Rad). The samples were run for 2 hours at 100 volts in running buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH 8.3). The proteins were transferred onto an Immun-blot PVDF membrane (Bio-Rad) for 2 hours at 4°C in transfer buffer (25 mM Tris, 192 mM Glycine, 20% v/v methanol, pH 8.3). The membranes were blocked for 1 hour with 5 % non-fat milk (Carnation) in 1X Tris buffered saline/Tween 20 (20 mM Tris, 500 mM NaCl, 0.05% Tween 20). Thereafter, it was incubated with Rep antibody (1F) diluted 1: 5000 in 1X TTBS, washed three times with 1X TTBS for 10 minutes each, and then incubated with Goat anti-mouse HRP-conjugated secondary antibody(Biorad) diluted 1:20000 in 1X TTBS for 1 hour. Subsequently, the membrane was washed three times in TTBS for 10 minutes, and incubated in 12 mLs of Immun-Star HRP Chemiluminescent reagent (Bio-Rad) for 5 minutes. The membrane was wrapped in cling film and exposed to film for the appropriate time for good signal development.

**P5-RepShuttle Virus Production**

Twenty 15 cm dishes containing HEK 293 cells were cultured in Dulbecco's Modified Eagle's Medium supplemented with 5% fetal bovine serum and 100 U/mL Penicillin and streptomycin. The cells were triple transfected with the following constructs (see Appendix A for vector maps): P5RepShuttle vector, rep-cap helper plasmid pAAV-RC, and adenovirus helper

plasmid pAdhelper, using Polyethylenimine (1 mg/ml). The mixture was vortexed and incubated at room temperature for 5 minutes before being applied to the cells. The transfected cells were incubated at 37°C for 72 hours, at which time they were then harvested and lysed by three freeze/thaw cycles. The crude lysate was clarified by centrifugation and loaded onto two discontinuous iodixanol step gradients by gently underlaying the crude lysate with 15% (9 ml), 25% (6 ml), 40%, and 60% (5 ml) iodixanol (OptiPrep). The gradients were centrifuged at 70,000 rpm for 1 hour. Five milliliters of the 60–40% interface was removed and dialyzed against PBS or loaded onto a heparin column. An empty Bio-Rad column was loaded with 3 µl of heparin-agarose (Sigma H-6508) and washed with 20 mLs of 1xTD (137 mM NaCl, 15 mM KCl, 10 mM $Na_2PO_4$, 5 mM $MgCl_2$, 2 mM $KH_2PO_4$, pH 7.4). The column was loaded with the pulled iodixanol fraction and subsequently washed with 20 mLs of 1xTD. The bound virus was then eluted with 1xTD plus 0.4M NaCl. Aliquots of the virus were stored at -80°C for future use. The titer was determined to be $1.6 \times 10^{12}$ viral particles per ml.

**DNA Walking Speedup PCR**

The DNA Walking Speed Up PCR kit (Seegene) employs a proprietary Anneal Control Primer (ACP) technology, which enhances the specificity of primer binding and PCR product formation. The kit is used to amplify unknown sequences flanking known sequences. The PCR strategy (shown in Figure 4-2) is as follows: One of the four provided ACP primers (ACP1, ACP2, ACP3, ACP4) and the designed target (vector) specific primer 1 (TSP1) are used in the first PCR reaction. The second PCR, the first nested PCR, uses an ACPN primer and second target specific primer (TSP2) to amplify from the 1st PCR product. The last PCR uses a universal primer and target specific primer 3 (TSP3) to amplify from the second PCR product. The PCR reaction for the 1st reaction was as follows: 94°C (5 min), 42°C (1 min), 72°C (2 min), 30 cycles of 94°C (30 sec), 55 °C (30 sec), 72°C (100 sec), and a final step at 72 °C (7 minutes). The

second PCR reaction conditions: 94°C (3 min), 30 cycles of 94°C (30 sec), 55-60°C (30 sec), 72°C (100 sec), and a final step at 72°C (7 minutes). The last PCR reaction conditions were the same as the second PCR conditions, except the annealing step was 60-65°C (30 sec). Five to ten microliters of the final PCR product were loaded on a 1% agarose gel for inspection or southern analysis. In addition, the bands were excised from the gel and individually cloned into a TA cloning vector for sequencing. The sequence information for the target specific primers are as follows: 5'ACCACCTTCAGTGCGGCAAA3'(TSP1),

5'ACAGTACTCCACGGGACACGGTCAG3'(TSP2),

5'ATTCAGAGCCTCGCCCCATTGG3'(TSP3).

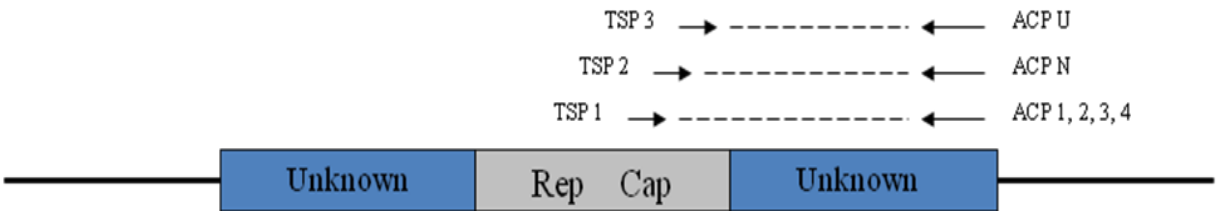



Figure 4-2. Nested PCR scheme for DNA-Walking PCR. The first PCR reaction involved primers TSP1 and ACP1 or ACP2, 3, 4. The second PCR reaction (1st nested PCR) involves the use of TSP 2 and ACPN primers. The third PCR reaction (2nd nested PCR) uses TSP 3 and ACP U primers. Note that the TSP1, 2, 3 are primers designed for the right end of the AAV 2 vector in the Cap sequence as marked.

**Junction Assay: Detecting AAVS1 Integration by PCR-Southern.**

A hybrid PCR-southern blot approach was used to determine if the vectors used in the experiments are capable of integrating into AAVS1. In this approach, PCR was used to amplify AAV-AAVS1 junction sequences from genomic DNA isolated from total cells infected or tranfected with the P5RepShuttle vector or wild type AAV 2 vector. The AAV ITRs are commonly detected at the AAV-AAVS1 junction sites and there is no well defined break point, but clusters of break points have been identified near the RBS of the ITR. Therefore, the following primers were used to allow isolation of junction sequences by PCR: the forward

primer was specific for the right AAV ITR (ITR1: 5'-aggaacccctagtgatggag-3') and the reverse primer falls on AAVS1 near a BamHI restriction site on the host sequence (AAVS1dRBS: 5'-caccacgtgatgtcctctga-3'). The PCR scheme is shown in Figure 4-3.

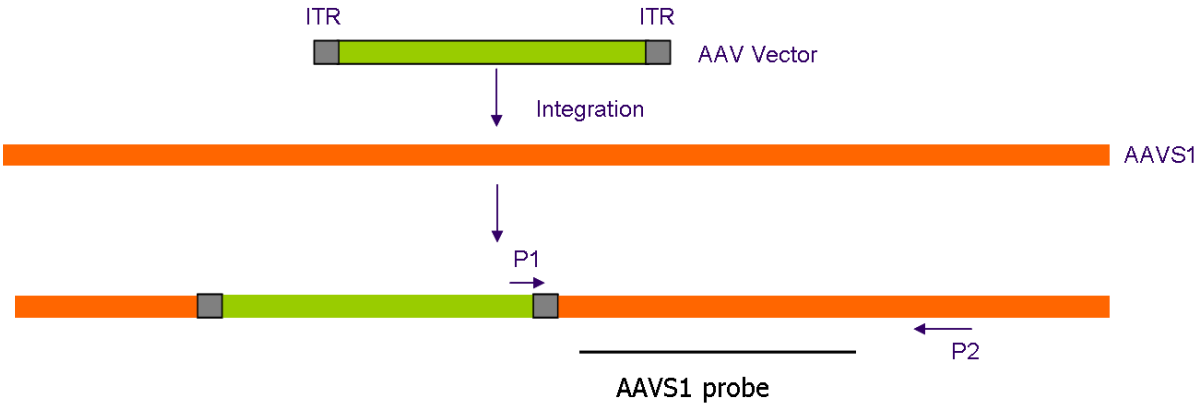



Figure 4-3. PCR scheme for detecting AAV-AAVS1 junctions. The figure represents a theoretical integration event of an AAV vector into AAVS1. To detect AAV-AAVS1 junctions one primer near the right ITR and a second primer in the target site is used in a PCR reaction. The PCR products are then hybridized with an AAVS1 probe designed to detect those products. Note that this system is only capable of detecting right end AAV-AAVS1 junctions.

The PCR conditions were set up using HotStarTaq master mix (Qiagen) as follows: an initial heating at 94°C for 15 min (hot start) followed by 30 cycles at 94°C for 1 min, 57°C for 1 min, 72°C for 2 min, and a final elongation at 72°C for 10 min. The PCR products were separated on a 1% agarose gel and then transferred onto a positive nylon filter membrane (Hybond N+, Amersham). The filter was hybridized with a 441 bp AAVS1 probe generated using the following primer combination (AAVS1p1: 5' accctatgctgacaccccgt3'; AAVS1p2: 5'cgcagaagccagtagagct3') using non-infected HeLa DNA as the template.

**DIG-Labeling of AAVS1 probe**

PCR DIG Probe Synthesis Kit (Roche) was used to generate and label a 441 bp AAVS1 fragment. 50-100 nanograms of HeLa genomic DNA was mixed with the following reagents:

10x PCR buffer (5µl), PCR DIG labeling mix (5µ l), 10 µm forward and reverse primers (2 µl each), enzyme mix (1 µl), and water (up to 50 µl). The PCR conditions were as follows: initial denaturation (95°C for 2 minutes), 30 cycles of denaturation (95°C for 20 seconds), annealing (55°C for 30 seconds), elongation (72°C for 2 min), and a final elongation step (72°C for 7 minutes). The labeling of the probe was assessed by running a portion of the PCR product on an agarose gel. Both labeled and unlabelled PCR products were assessed this way. Successful labeling results in the PCR product migrating slower, hence higher in size, compared to the unlabelled control. The probe concentration used in hybridization was 2 µl of probe per ml of hybridization solution.

**Hybridization of PCR Products**

Hybridization of the PCR products was performed using DIG Easy Hyb (Roche) overnight at 42°C. Post hybridization the nylon filter was washed twice in 2XSSC (0.3 M sodium chloride, 0.03 M sodium citrate, pH 7.0), 0.1% SDS (sodium dodecyl sulfate) at room temp for 10 minutes each followed by washing in 0.5XSSC, 0.1%SDS at 68°C for 30 minutes. After hybridization and stringency washes the blot was incubated in washing buffer for 2 minutes, blocking solution for 30 minutes, antibody solution for 30 minutes, washed twice for 15 minutes in washing buffer, and finally incubated in detection buffer for 2 minutes. The membrane was incubated with CSPD for a few minutes, incubated for 10 minutes at 37°C. The membrane was wrapped in saran and exposed to film for the appropriate time period for signal development (5-15 minutes). The wash and block solutions, CSPD, and antibodies were all purchased from Roche. The antibody, Anti-Digoxigenin Fab fragments (Roche), was used at a dilution of 1:10,000 in blocking buffer.

**Rescue of Integrated Shuttle Vector**

To rescue integrated P5RepShuttle vector, 20 µg of genomic DNA, from transfected or infected cells, was digested with EcoRV, which cleaves upstream of the Amp/Ori and throughout

the cellular genome. The DNA was extracted with phenol/chloroform and precipitated with ethanol. The pellet was resuspended in 200 µl of water and ligated using 5 µl of T4 DNA ligase (NEB, 400 U/ul) overnight at 15 °C, to promote intramolecular circularization. The ligated DNA was precipitated and resuspended in 5 µl of water and transformed into SURE (Stop Unwanted Rearrangement Events) bacterial cells by electroporation. The transformed bacteria were grown on agar containing 100 µg/ml ampicillin.

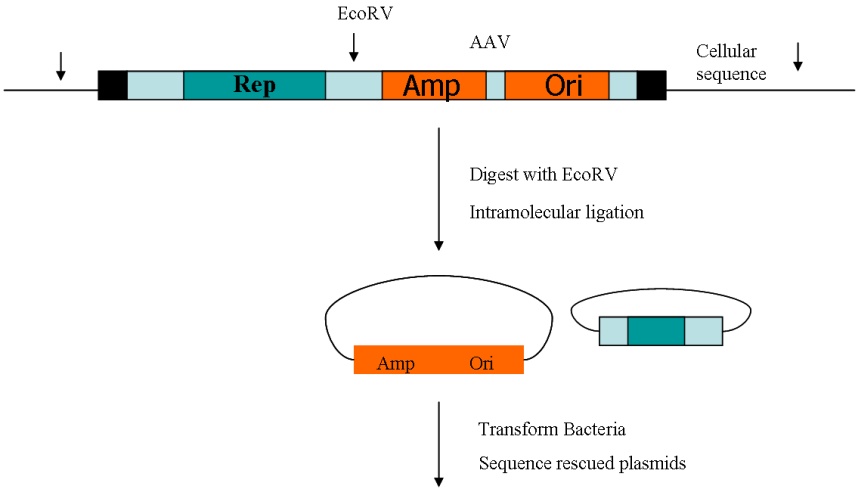



Figure 4-4. Shuttle rescue scheme to isolate provirus as plasmids. Genomic DNA containing integrated shuttle vector is digested with EcoRV which cuts upstream of the Amp and Ori sequences and throughout the cellular genome. The digested fragments are ligated under conditions that promote intramolecular circularization. These circularized fragments are transformed into bacteria and grown as plasmids for sequencing. Note that only those fragments containing both the Amp and Ori will be amplified and selected in bacteria.

Bacterial plasmids from the surviving bacterial colonies were isolated and sequenced using a primer in the AAV Ori near the right ITR (5'ACACCGAACTGAGATACCTACAGCG 3'). Sequencing was performed by the Interdisciplinary Center for Biotechnology Research (ICBR) facility (Cancer/Genetics Research Complex, University of Florida). The sequenced information was blasted against the NCBI human genome database to find matches for the rescued flanking cellular sequence to identify where the shuttle vector had integrated.

**Transformation of Competent Cells**

One shot TOP10 Electrocompetent cells (Invitrogen) were thawed on ice. One to two microliters of cloned DNA was mixed gently with the bacterial cells. The mixture was transferred to a chilled 0.2 cm electroporation cuvette and transformed according to manufacturer's instructions. The electroporated mixture was removed and placed in 250 µl of S.O.C. medium and shaken at 37°C for 1 hour. Ten to 200 µls of each mixture was plated onto pre-warmed ampicillin plates and incubated overnight at 37 °C.

SURE electroporation-competent cells (Stratagene) were transformed according to manufacturer's instructions. The SURE stands for Stop Unwanted Rearrangement Events. The cells were used to transform AAV vectors which contain the ITRs and for junction rescue of the shuttle vector, and because they provide high transformation efficiencies (10^10 cfu per microgram)

**Cloning of PCR products into TOPO TA Vector**

TOPO TA cloning Kit for sequencing (Invitrogen) was used to clone PCR products for sequencing. Briefly, 0.5 to 4 µl of PCR product was mixed with 1 µl of salt solution (1.2 M NaCl, 0.06 m $MgCl_2$), 1 µl of TOPO vector, and water up to 6 µl. The mixture was incubated for 5 minutes at room temperature and then placed on ice. One shot TOP10 electrocompetent cells were transformed with the mixture as previously described.

## Results

**Limit of Detection for Site-Specific Integration Using the Junction Assay.**

A HeLa cell line latently infected with AAV (Clone #2) was used to demonstrate detection limits of the assay. Decreasing amounts of genomic DNA from clone 2, generated by using 10 fold dilutions in water, was mixed with 640 nanograms of background uninfected Hela DNA for PCR reactions. The PCR products were probed with AAVS1 using Southern hybridization.

HeLa cells typically contain 6-10 pg of DNA per cell. Assuming a DNA content in HeLa cells of around 8 pg of DNA per cell, $1X10^5$ cells of Clone #2 would give around 800 ng of DNA (undiluted DNA sample in figure).

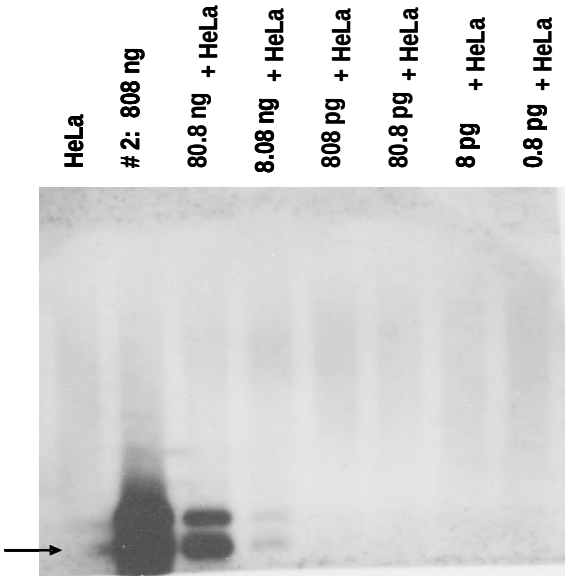



Figure 4-5: Southern blot to assess detection limit of junction assay. PCR product of genomic DNA latenly infected with AAV 2 was hybridized using an AAVS1 probe. Decreasing amounts of Clone #2 genomic DNA (using 10 fold dilutions) was used with 640 ng of uninfected HeLa background DNA for the junction PCR. Arrow marks the 0.5 kb size

The result (Figure 4-5) indicates that the detectable limit for AAV-AAVS1 junction sequences is 8 ng, which is equivalent to approximately 1000 cells in $10^5$ cells, or 1%. The control, uninfected HeLa cells, is negative for such AAV-AAVS1 junction, demonstrating the specificity of the assay. Overall, this information suggests that if 1% of infected cells contain integrated virus and form AAV-AAVS1 junctions within my primer range, I would be able to detect it using the junction assay. Therefore, this approach was used as a preliminary assessment to determine if the constructed shuttle vector was capable of site-specific integration.

**P5RepShuttle Vector Mimics Wild Type AAV With Regards to Rep expression.**

To assess if the shuttle vector was capable of expressing Rep, HeLa cells were either mock transfected or transfected with 3 µg of the following plasmids: pAV2, pSVAV2, P5RepShuttle vector, or P5RepShuttle vector plus pXX680 plasmid (4.5µg). pAV2 plasmid contains the wild type AAV2 genome. pSVAV2 is a construct expressing Rep under a constitutive SV40 promoter. pXX680 is an adenovirus helper plasmid containing the adenovirus genes necessary for AAV replication and expression.

Protein was extract from the transfected samples 48 hours post-infection for western blot analysis. The result (Figure 4-6) indicates that the P5RepShuttle vector expressed the Rep78 and 52 proteins, only in the presence of ad helper genes. Like wild type AAV, Rep is not detectable in the absence of ad helper genes. Although undetectable in the absence of adenovirus genes, enough Rep is presumably made early on allowing site specific integration to take place. It is important to note that Rep negatively autoregulates its expression by binding to its P5 promoter; the construct pSVAV2 expresses all the Rep proteins without such negative regulation, serving as a positive control.

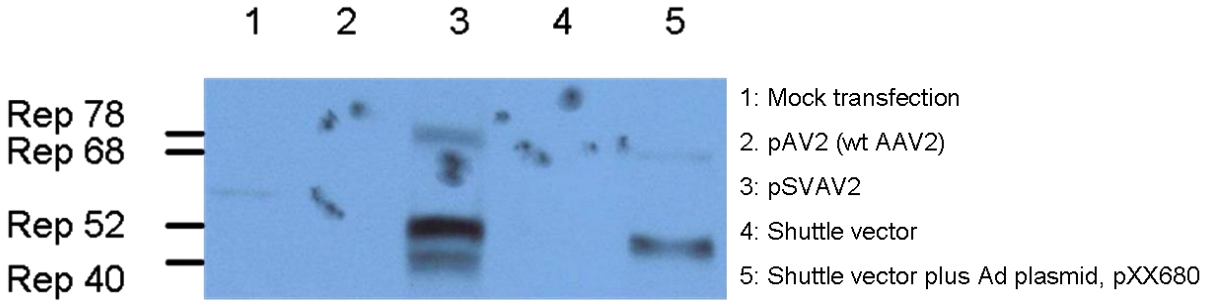



Figure 4-6: Western blot for Rep expression from the P5RepShuttle construct.

**Transfected P5Rep Shuttle Integrates into AAVS1**

In the absence of Rep expression, the shuttle vector would not be targeted to AASV1. To verify whether the P5-Rep shuttle vector is capable of expressing enough Rep and integrating

site-specifically, HeLa cells were transfected with 1.5 micrograms of P5RepShuttle plasmid using PolyFect transfection reagent (Qiagen). Forty-eight hours post-tranfection the cells were expanded from a 6 well dish to a T25 flask. A day later, the confluent cells were split 1:10 into 2 T25 flasks. The rest of the cells were harvested (the 3 days post-transfection sample point). Every four days, cells from the two T25 flasks were mixed and split into two new T25 flasks at 1:10. The rest of the cells were collected for junction analysis. The collection days used in the junction PCR are listed in the figure. One microgram of DNA from each sample time point was used in a junction PCR. The PCR products were probed for AAVS1 using southern hybridization.

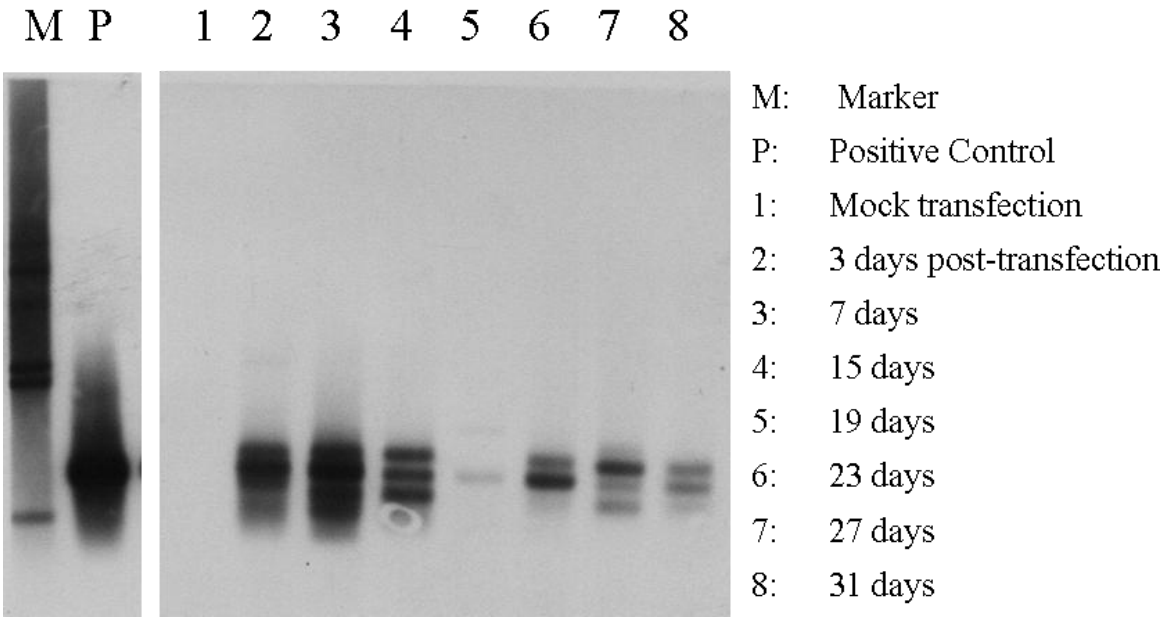



Figure 4-7: Southern on PCR products generated from HeLa genomic DNA transfected with P5Repshuttle vector at various days post-tranfection. M indicates the DNA ladder marker for size identification. P indicates the positive control sample used to verify success of PCR reaction. The numbers indicate the different samples.

The passaging of the cells was used to dilute out the number of transfected plasmids in the cells which could inhibit the PCR reaction by saturating the primer binding sites. The result (Figure 4-7) indicates that the P5-RepShuttle vector was capable of integrating into the

Chromsome 19 target site. Moreover, integration can be detected, as judged by persistence of AAV-AAVS1 junction sequences, even at late passage (approximately 31 days post-tranfection). In addition, this suggests that enough Rep is expressed to allow targeting of the plasmid to AAVS1, although undetectable by western blot.

**Infected P5Rep Shuttle Integrates into AAVS1**

The previous data indicated that the transfected shuttle vector plasmid was capable of integrating, due to sufficient Rep expression. Therefore, the next objective was to package the P5RepShuttle vector into AAV particles and assess whether the P5RepShuttle virus was capable of infecting HeLa cells and site-specifically integrating.

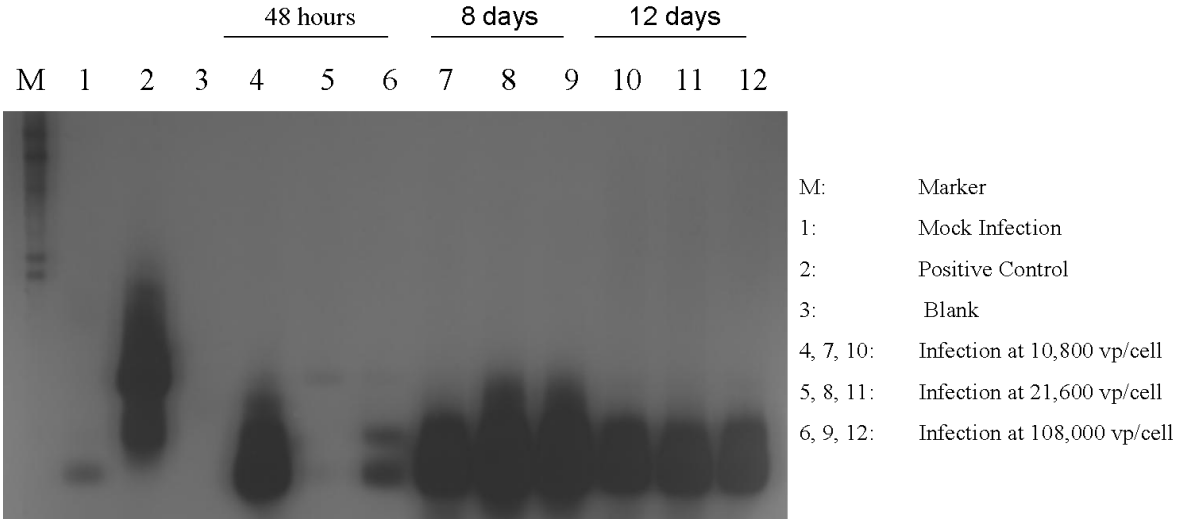



Figure 4-8: Southern blot on PCR products from HeLa cells infected with P5RepShuttle vector at various infection doses.

HeLa cells were infected with various multiplicities of infection (viral particles/cell) of P5RepShuttle virus: 10,800 or 21,600 or 108,000 vp/cell. Forty-eight hours post-infection the cells were split at 1:10 into new tissue culture plates. The rest of the cells were collected for PCR (shown as the 48 hour time point sample). The other time points included 8 days post-infection and 12 days post-infection. One microgram of genomic DNA from each sample was used in the junction PCR as described previously. The results (Figure 4-8) indicate that the

shuttle vector virus was capable of infecting HeLa cells and integrating into AAVS1. Moreover, the increase in dose of viral infection didn't increase the intensity of junction product suggesting that maximal integration took place even at the lowest dose tested, limited only by the intrinsic efficiency of integration. In addition, similar to the transfection experiment, junction product was detectable even at the latest harvesting time point (12 days post infection).

**Shuttle Vector Rescue Isolates AAV-AAVS1 Junctions, But Not Those from Other Sites**

The previous data indicated that the shuttle vector was capable of integrating into the targeted site. To determine if the vector also integrated into other sites, genomic DNA from both transfected and infected HeLa samples were used for junction rescue. The rescue of the integrated virus isolated only AAV-AAVS1 junctions and not other AAV-cellular junctions, even after numerous attempts. The rescued products were sequenced to identify the precise junction sequences. Three junctions (Table 4-1) were unambiguously mapped to Chromosome 19 site in AAVS1, at different locations (based on UCSC Genome Browser numbering).

Table 4-1: P5Shuttle Vector-Chromosome 19 junction sequences isolated through rescue approach.

| AAV Sequence | Chr. 19 Sequence | Location |
|---|---|---|
| CGACTCCACCCCTCCAGGAACCCCTAGTGA | TGGCCCAGATCCTTCCCTGCCGCCTCCTTCAG.. | 60319669 |
| GCTCGCTCGCTCACTGAGGCCGGGCGACCA | GCCCCCGAGTGCCCTTGCTG.. | 60319887 |
| CCTCTCTGCGCGCTCGCTCGCTCACTGAGGC | CGCCTTTCAGGGGGACCCAGGGCACCAGAACTCCC | 60298701 |

**AAV-Chromosome 3 and AAV-Chromosome 6 Junctions Isolated Using DNA Walking PCR**

Since the P5Shuttle vector experiments failed to isolate other AAV-cellular junctions, an alternative strategy was used in hopes of identifying other integration sites. To do this, HeLa cells were infected with wild type AAV 2 at 10,000 viral particles per cell. The wild type AAV 2 virus was produced in HEK 293 cells using AAV 2 plasmid pAV2 and helper vector pDG. Wild type AAV virus was used since it is unclear whether the packaged shuttle virus was fully

functional. The DNA walking PCR system by Seegene was used to isolate cellular sequences flanking AAV. A small aliquot of the final product was used in AAVS1 southern hybridization to see if this system could pickup AAV-AAVS1 junctions. Hybridization using AAVS1 and AAV probes revealed that the DNA walking PCR kit was able to isolate AAV-AAVS1 junction, but its relative presence is less than the other PCR products, based on the intensity of the PCR products (Figure 4-9).

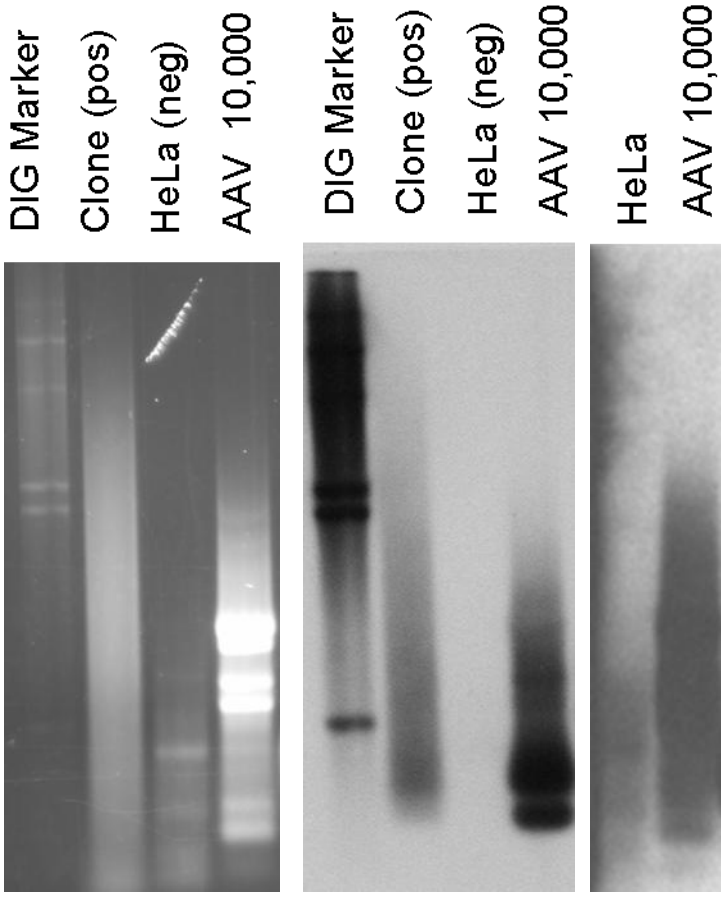



Figure 4-9: Southern blot on PCR products generated by DNA walking Speedup PCR system. A gel picture of the PCR products (left) indicate strong signals for HeLa cells infected with wild type AAV2 at 10,000 vp/cell. HeLa (negative) is mock infected DNA used in PCR reactions. Clone (pos) represents sample used as positive control to determine if the PCR system can amplify AAV-AAVS1 junction sequences. (Middle) AAVS1 hybrdization on PCR products. (Right) AAV hybridization on PCR products.

To identify what those PCR products represented, they were cloned into the TOPO TA vector (Invitrogen) for sequencing. Two junctions were unambiguously mapped to chromosomes 3 and 6. Another junction mapped to several chromosomal sites due to repetitive sequences at the junction making it difficult to identify the precise integration site. Numerous additional attempts to isolate more junctions resulted in failure. The junctions sequences identified are presented below in Table 4-2.

Table 4-2: AAV2-Chromosomal junction sequences identified using the DNA Walking Speed Up PCR kit.

| AAV Sequence | Chromosome Sequence | Location |
|---|---|---|
| ACAGTACTCCACGGGACACGG | TCAGTGCTAGAGGACAGCTTCAACTCCC CATAATTTCATCTCTGACC | Chr. 3 |
| CGCGCTCGCTCGCTCACTGAGGCCGGGCGA CCAAAGGT | GGTAGAAAAGGTATATCCTCGTAGAAAA ACTAGA | Chr. 6 |
| TGGAGTTGGCCACTCCCTCTCTGCGCGCTC | GCTCTGGAGACCTGATGCTG GGGAAGGGCA TGCCTGGCATCACCACACACCTGGGGGGAG ACAGGAGCCTGGGGCCGGTGGGCCCACACA | Chr.1, Chr. 15, Chr. 16. |

## Discussion and Limitation of Study

In an attempt to identify the distribution of AAV integrants in the presence of Rep, two approaches were used: the first approach involved the use of a Rep-containing shuttle vector to isolate integrated vectors as plasmids in bacteria, and the second approach involved the use a commercial kit to PCR amplify AAV-cellular junctions.

The first obstacle has been that the P5RepShuttle didn't work as efficiently as hoped. It was a novel approach which has the potential to identify integration sites. Initially, these experiments were hindered by the difficulty in packaging the shuttle vector, because there was significant homology between the shuttle vector and the helper constructs (i.e the Rep, Amp, and Ori sequences were all present in helper vectors). Any recombination among these vectors during

the replication and packaging process could have altered the Amp and Ori of the packaged vector, thus affecting the subsequent infection in Hela cells and the corresponding rescue experiments. In an attempt to minimize this, an approach was used to reduce this homology, but it resulted in low titers of the shuttle vector, not sufficient for cell culture experiments.

Despite these difficulties, several P5RepShuttle-AAVS1 junctions were isolated using the rescue approach, suggesting that some of the virus was functional. The lack of rescue of other junctions can be explained as follows: (i) larger circularized fragment (>8 kb), due to integration of AAV as possible concatemers, may not transform efficiently in bacteria , making the isolation of these junctions difficult, (ii) the choice of restriction enzyme greatly limits the efficiency of the approach, since an enzyme must not cut in the Amp and Ori for efficient rescue, (iii) integrants in chromosome 19 are more likely to integrate largely intact, compared to other random integration sites, resulting in predominantly AAV-AAVS1 junctions being isolated, (iv) the large homology (Rep, Amp, Ori) among the shuttle construct and helper vectors may have contributed to packaging of disrupted Amp and Ori sequences affecting integration and rescue, (v) the imprecise nature of integration could have result in many integration events, but most of these possibly suffer from partial or complete loss of the necessary Amp and Ori sequences required for rescue, making impossible the isolation of these junctions using the rescue approach, and (vi) it is also important to note that the integration of the P5RepShuttle vector was allowed to occur in the absence of selective pressure, possibly resulting in a limited number of integration events taking place.

The DNA Walking PCR kit was used as an alternative approach to try and identify integration sites not possible by the shuttle approach. To simply the interpretation of the results, wild type AAV2 was used. The manufacturer claims that this approach is superior to many other

complicated methods such as inverse PCR, and ligation mediated PCR, therefore it was used instead of those other approaches. Interestingly, Southerns on the PCR products generated using this system were able to identify AAV-AAVS1 junctions, along with other PCR products not positive for AAVS1, presumably representing other AAV-cellular junctions. However, this system suffers from primer bias and amplification conditions can greatly affect results (data not shown). Sequencing of cloned PCR products unambiguously identified Chromosomes 3 and 6 as AAV integration sites, but chromosomes 1, 15, and 16 may also be other sites of integration. Comparison of these sequences in detailed didn't reveal a mechanism of integration, because there wasn't any significant homology between the vector and the chromosomal site. Other sequenced junctions (not described) were complex and could not be explained as simple AAV-chromosomal junctions. Some of these represented AAV-AAV and AAV-Adenovirus plasmid junctions. A potential obstacle in isolating and identifying junctions is that the PCR approaches used in junction detection and sequencing can be a hindered by the difficulty for the DNA polymerase to amplify through the ITRs, since the presence of any secondary structure of the ITRs at the junctions can cause the stalling of the DNA polymerase. In summary, the small scale of the results precludes any statements about whether Rep can minimize random integration.

Future research should continue these experiments to focus on studying whether Rep can minimize random integration and provide stable, safe integration. The limited results using the described approaches demand innovative approaches for studying AAV integration. One suggestion would be to use a novel shuttle vector containing a selection cassette that can be used for selection of both mammalian and bacteria cells. This new shuttle vector can then be used in a co-infection with a Rep vector to target the integration of the shuttle vector to AAVS1. The distribution of integration of the shuttle vector can then be assessed in the presence or absence of

Rep expression in total selected cells or in individual cell clones. In addition, by providing the Rep as a separate vector, the issue of rearrangements of the shuttle vector, due to large regions of homology with helper vectors can be avoided, potentially making junction isolation and analysis easier.

A recent report has indicated that a great deal of variability exists with regard to integration events with wild type AAV within 48 hours post-infection (Drew et al., 2007). This observation remains to be validated using alternative approaches, and therefore, it is important to characterize the distribution of integration for wild type AAV and/or Rep containing AAV vectors, since Rep is an important viral protein required for targeted integration. Elucidating the details would provide renewed interest in the development of AAV as an integrating vector for certain gene therapy applications. Moreover, from a virology perspective, it would greatly contribute to our understanding of how Rep can function to regulate the life cycle of AAV, priming it for integration in the absence of helper virus.

# CHAPTER 5
# DETERMINE THE ROLE OF CELLULAR DNA REPAIR PROTEINS ON SITE-SPECIFIC INTEGRATION

## Introduction

Very little is known about how cellular proteins interact with the AAV genome, processing it for AAV integration or affecting its tranduction. Many cellular proteins have been identified that interact with the AAV Rep protein and were shown to modulate AAV DNA replication (Nash et al., 2009). Interestingly, amongst these identified Rep-interaction proteins, several belong to the non-homologous recombination or other DNA repair pathways, including Ku 70/Ku 80, DNA-dependent protein kinase, Rad50, and PARP1. These same proteins may have a role in AAV integration, but these studies have not been actively pursued by others.

Many studies have indicated that upon AAV infection, the host cell detects the AAV single stranded DNA as damaged DNA (Jurvansuu et al., 2005; Cervelli et al., 2008), suggesting that host repair proteins can passively and/or actively compete for AAV binding. In addition, Zentilin et al., used quantitative chromatin immunoprecipitation to demonstrate that both Ku86 and Rad52 can bind the AAV ITRs, processing the AAV genome via alternate pathways that can affect transduction expression (Zentilin et al,, 2001). In the absence of Ku86 they observed greater transduction, whereas, in the absence of Rad52, rAAV transduction was greatly inhibited. This study suggests that there is competition between non-homologous (i.e Ku86) and homologous repair proteins (i.e Rad52) in processing AAV expression. It remains to be seen whether these proteins have other roles, such as processing wild type AAV infection for site-specific integration.

There have been only three reports to date that suggest a direct role of the cellular repair machinery in integration. Sanlioglu et al., used ATM deficient and proficient cells to demonstrate that rAAV integration is elevated in ATM deficient cells, observing a 7 fold increase in the

number of GFP positive clones in ATM deficient cells compared to controls (Sanlioglu, S., Benson, P.K., and Engelhardt, J.F., et al., 2000). More interestingly, Song et al. infected both C56/BL6 and SCID mice hepatocytes with rAAV vectors and used partial hepatectomy to dilute out the episomal AAV molecules to measure integration, observing transgene expression in 10% of hepatocytes in C56/BL6 mice 8 weeks post-hepatectomy, whereas, over 40% of hepatocytes in SCID mice expressed the transgene at the same time point. SCID mice suffer from severe combined immunodeficiency disorder, because they are lacking expression of DNA-dependent protein kinase catalytic subunit (DNAPKcs), suggesting that DNAPKcs is an inhibitor of both rAAV and wild type AAV integration, the latter determined in an *vitro* assay (Song et al., 2004). Last but not least, Yamamoto et al., identified a new role for TAR RNA loop binding protein 185 (TRP-185) in AAV 2 site-specific integration, suggesting that TRP-185 inhibits wild type AAV integration near the AAVS1 RBE by enhancing integration more downstream into AAVS1 (Yamamoto et al., 2007). They postulated that TRP-185 acts to enhance Rep 68 3' -5' helicase activity by serving as a molecular chaperone. Taking into account these observations, it is a bit confounding how cellular repair proteins act on AAV genome processing for it for integration, whether random or site-specific (AAVS1).

To contribute knowledge to this area of research, my specific aim was to assess first and foremost, whether non-homologous end joining has a direct role in AAV site-specific integration. To do this, glioblastoma cells lines, M059J and M059K, were used in co-infection experiments and clones were analyzed by Southern hybridizations to assess site-specific integration. The M059J cells lack expression of DNAPKcs and as a result are defective for non-homologous repair. The M059K cells express functional DNAPKcs and are competent for non-homologous repair.

## Experimental Design and Methods

### Cell Lines

M059J and M059K cell lines were purchased from American Type Culture Collection (Manassas, VA). The M059K cells are resistant to ionizing radiation damage, capable of fixing the damage. The M059J cells are 30 fold more sensitive to ionizing radiation than M059K cells. M059K cells express functional DNA-dependent protein kinase catalytic subunit (DNAPKcs), whereas M059J cells do not express DNAPKcs. These cells were maintained in a 1:1 mixture of Dulbecco's modified eagles medium and Ham's F12 medium supplemented with 0.05 mM non-essential amino acids, 10% fetal bovine serium, and 100 U/mL of penicillin and streptomycin. They were maintained in a 37 °C humidified incubator with 5% $CO_2$.

A human cervical cell line (HeLa) was maintained in Dulbecco's Modified Eagles Medium (GIBCO) supplemented with 10% fetal bovine serum and 100 U/mL of penicillin and streptomycin. Cells were maintained in a 37°C humidified incubator with 5% $CO_2$.

The Ligase I (GM16097) and Ligase IV (GM16089) cell lines were purchased from the Coriell Institute (Camden, NJ) and maintained in Minimum Essential Medium supplemented with 2 mM L-glutamine, 100 U/ml penicillin and streptomycin, and 10% fetal bovine serum. The Ligase I cell line is defective for ligase I function. The Ligase IV cell line is defective for ligase IV function. The ligase I enzyme is important in homologous repair and the ligase IV enzyme is important for non-homologous end joining repair.

### Construction of Recombinant AAV Vectors

P5UF11 was produced by inserting the P5 sequence into a Kpn1 site of UF11 vector (Vector Core, University of Florida). The P5 sequence was PCR amplified from wild type AAV 2 plasmid with primers containing internal Kpn1 sites. The orientation of the inserted P5 was examined by sequencing. P5UF11 contains the P5 sequence, CAG promoter (CMV enhancer and

chicken beta-actin promoter), GFP, and the Neomycin cassette flanked by the ITRs. The P5UF11 vector was packaged into AAV 2 capsid in HEK 293 cells via cotransfection with pAAVRC and pADhelper plasmids.

P5PGKHygroGFP was constructed as follows: pMSCVHyg (Clontech) was digested with XhoI and HindIII to isolate a 1872 bp fragment containing the PGK promoter and hygromycin coding sequence. This fragment was cloned into pdsP5AAV-CB-EGFP vector backbone isolated by digesting with XhoI and HindIII. This new construct, called dsP5PGKhygo, was digested with XhoI and SacI to isolate the PGK and hygromycin sequences and cloned into P5UF11 digested with XhoI and SacI. The vector was packaged into AAV 2 capsids using HEK 293 cells and helper vector pDG.

pSVAV2 was constructed from pAV2 and pSVRep (a gift from Dr. Falck-Pedersen). The pSVRep was digested with EcoRV and HindIII to isolate a fragment containing the SV40 promoter and Rep78 sequences. This fragment was inserted into pAV2 digested with BmgBI and HindIII.

pdsP5AAVNeoR: This self-complementary AAV vector was created by cloning the P5 and Neomycin sequences into pdsAAV-CB-EGFP (a kind gift from Dr. Arun Srivastava). The P5 sequence was amplified using P5UF11 and inserted into a Kpn1 site of pdsAAV-CB-EGFP, to generate a pdsP5AAV-CB-EGFP construct. The neomycin cassette was isolated from P5UF11 using XhoI and SacII and inserted into the backbone of pdsP5AAV-CB-EGFP digested with XhoI and SacII. The constructed pdsP5Neo was verified by restriction digestions. The vector was packaged into AAV 2 capsids using helper vector pACG2.

pdsAAV-SVRep78: This self-complementary AAV vector was created by isolating a 1982 bp fragment containing the SV40 and Rep78 sequences from pSVRep and inserting it into the

backbone of pdsAAV-CB-EGFP digested with Kpn1 and HincII. The constructed vector was verified by restriction digestions, and packaged into AAV 2 capsids using helper vector pACG2.

See Appendix A for vector maps.

**Cloning of Infected Cells**

Cells infected with the different viruses were selected for 2 weeks in the presence of Geneticin (GIBCO) (HeLa: 600 ug/ml; MO59J, M059K: 100 ug/ml) or Hygromycin B (GIBCO) (Ligase IV: 75 ug/ml). The resistant cells were counted and seeded at 1 cell per well into several 96 well plates, and subsequently expanded until they were confluent into T25 flasks or 100 mm dishes, at which time they were harvested for southern hybridization analysis.

**Southern Blot Analysis**

DNA from individual clones resistant to neomycin was isolated using DNAeasy tissue kit (Qiagen). 10-20 ug of DNA was digested with EcoRV and XbaI for 12-16 hours, and loaded onto a 0.8% agarose gel. The DNA was transferred from the gel onto a nylon filter membrane and hybridization was carried out at 65 degrees overnight with 32p labelled probes (AAVS1, Neo-GFP, or Neo). The DNA probes were radio-labeled using RadPrime DNA labeling kit (Invitrogen) following manufacturer's instructions. The AAVS1 probe consisted of a 3.5 kb fragment generated by digesting an AAVS1 plasmid (pRVK) with EcoRI and Kpn1. The Neo-GFP probe (2 kb) was generated by digesting P5UF11 with XbaI and BamHI. The 1 kb Neo probe (for self-complementary AAV vectors) was generated by digesting P5UF11 with XhoI and BamHI. The 1.9 kb Hygro probe was generated by digesting P5PGKHygro with XhoI and SacI. Hybridization was performed in Church Buffer (0.5 M sodium phosphate, 7% SDS, 1 mM EDTA) at 65°C overnight. The following day, the blots were washed in 2XSSC, 0.1% SDS (65°C for 15 minutes), 1X SSC, 0.1% SSC( 65°C for 15 minutes), 0.5xSSC, 0.1% SDS (65°C for 30 minutes) and final wash in 0.1X SSC, 0.1% SDS (65°C for 15 minutes). The nylon filter

was rinsed in 2XSSC, wrapped in Saran and exposed to film for appropriate times for signal development using autoradiography.

## Results

### More Junction Product Seen in M059K Cells Compared to M059J cells.

To assess site specific integration in these cells lines, they were infected with wild type AAV2 (pAV2) at different doses. AAV-AAVS1 junction formation was assessed at 48 hours post infection using the junction assay (Figure 5-1), described in material and methods on page 52. The results demonstrate that increasing the dose of wild type AAV2 did not increase the intensity of junction product in both J and K cells, suggesting that an intrinsic efficiency of integration is taking place in both cells. In addition, more junction product was detected in the M059K cells compared to the M059J cells.

### Time Course for Junction Formation in M059J and M059K Cells

To determine the kinetics of junction formation in these cells lines, junction product was assessed at 48 hours, 1 week, and 2 weeks post wild type AAV 2 infection (10^5 vp/cell).

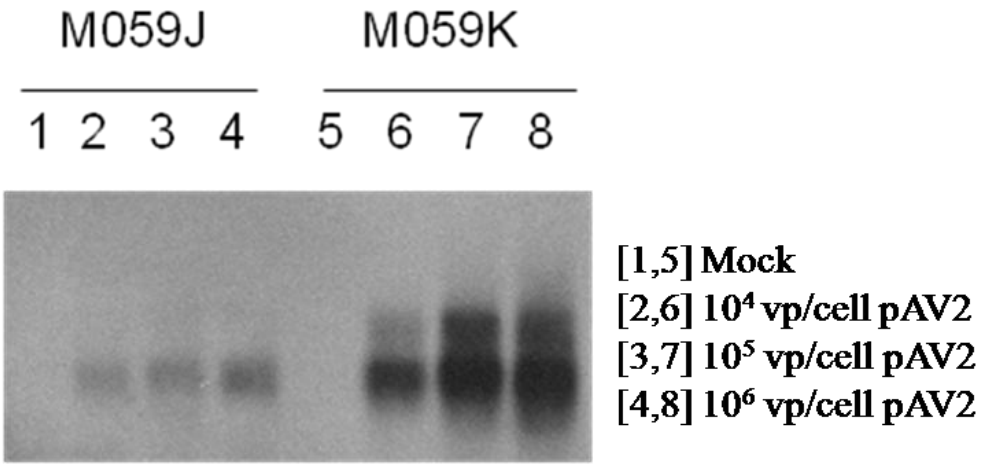



Figure 5-1. Junction product formation in M059J and M059K at different doses of wild type AAV 2 infection. The cell lines were infected with different dose (viral particles per cell) of wild type AAV2. Genomic DNA from infected cells was isolated 48 hours post-infection for junction assay.

The results (Figure 5-2) verified that J cells produce less junction product; however, by 1 week post infection the junction product in J cells increased. Interestingly, in the K cells the junction product also picked up by 1 week, and at all time points it was more intense compared to the J cells. This suggests that in the absence of DNAPkcs, AAV site specific integration is delayed.

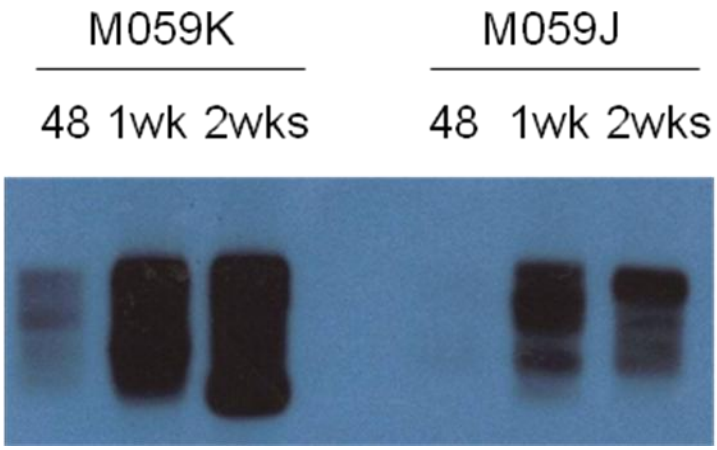



Figure 5-2: Time course of junction formation in M059J and M059K cells. Both cells were infected with 10^5 vp/cell. Genomic DNA was isolated at 48 hours, 1 week or 2 weeks post infection for junction assay.

An alternative infection approach was used to assess integration, which would allow for selection and cloning. M059J and M059 K cells were infected with P5UF11 and pSVAV2 at a ratio of 50:1, respectively, at a total dose of 10^6 vp/cell. The 50:1 ratio of these two single-stranded viruses was used because we found it provides optimal site-specific integration (Zhang et al., 2007). The infected cells were selected with Geneticin (G418) for 2 weeks in duplicates. Thereafter, the resistant colonies from one set were pooled for junction PCR to assess integration. Interestingly, upon selection, a significantly greater amount of junction product was detected in selected K cells, in contrast to the selected J cells (Figure 5-3). This suggested that more stable integration has taken place in K cells compared to J cells in the presence of selective pressure.

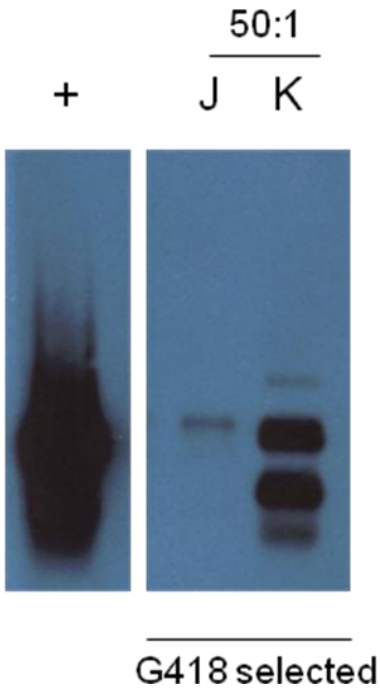



Figure 5-3. Junction Assay on M059J and M059K cells co-infected with P5UF11 and pSVAV2 at 10^6 vp/cell. Infected cells were selected for 2 weeks in G418 and surviving cells from each cell line was pooled for junction PCR. (+) represent a sample positive for AAV-AAVS1 junctions. The infected samples (J, K) are shown, as well as the ratio of the two viral vectors (50:1, P5UF11 to pSVAV2).

**DNAPKcs is an Inhibitor of Single Stranded DNA Site-Specific Integration**

To get a more quantitative assessment, the colonies from the second set of selection were expanded into larger tissue culture dishes and grown in the presence of G418. Genomic DNA was isolated from nineteen individual colonies cloned from both M059J and M059K infected cells for southern analysis. Interestingly, the data revealed something completely different than what was suggested by the junction assay: (i) more disruptions of the AAVS1 site occurred in M059J cells and 87.5% of clones had the neomycin vector integrated into AAVS1 (Figure 5-4), and (ii) in contrast, less disruption of AAVS1 was seen in M059K cells with only 60% of clones having the AAV neomycin cassette integrated into AAVS1 (Figure 5-5). The data are summarized in Table 5-1. The percentages of random and site-specific integration may be

slightly overestimated, since 4 M059K and 3 M059J clones did not have detectable neomycin bands, and presumably in these clones AAV is integrated randomly because these clones were grown in the presence of G418. Therefore, if this assumption is correct, then the M059K cells display 47% AAVS1 specificity, whereas, the M059J cells have 73% specificity. Regardless, in the absence of DNAPKcs (i.e M059J cells) there is significantly more specificity of integration.

Table 5-1. Summary of M059K and M059J single-stranded infected clones analyzed by southern hybridizations.

| | M059K Clones | M059J Clones |
|---|---|---|
| Number of clones disrupted for AAVS1: | 9 | 15 |
| Number of clone with detectable Neo: | 15 | 16 |
| Percentage of disrupted clones with Neo co-linked | 9/9 (100%) | 14/15 (93%) |
| AAVS1 specificity: | 9/15 (60%) | 14/16 (87.5%) |
| Random: | 6/15 (40%) | 2/16 (12.5%) |

**Self-Complementary AAV Vectors Integrate Site-Specifically in HeLa Cells**

DNAPKcs has been shown to have a role in concatemer formation of input linear ssDNA genome and the circularization of input self-complementary DNA genome. Of interest would be to determine if self-complementary vectors show different integration properties in the absence or presence of DNAPkcs. However, it has not been shown if scAAV vectors are capable of site-specific integration. To determine if scAAV vectors are capable of site-specific integration, HeLa cells were co-infected with two self complementary vectors (dsP5AAVNeoR and dsAAV-SVRep78) at different ratios and infection doses (viral genomes per cell). The integrating scAAV, dsP5AAVNeoR, was used for targeted integration ito AAVS1, and the helper vector, dsAAV-SVRep78, was used to provide the Rep protein for targeting the dsP5Neo vector to AAVS1. Different ratios of the two vectors were used to decide which ratio provided optimal integration detectable by the junction assay analyzed 1 week post-infection (Figure 5-6 A).

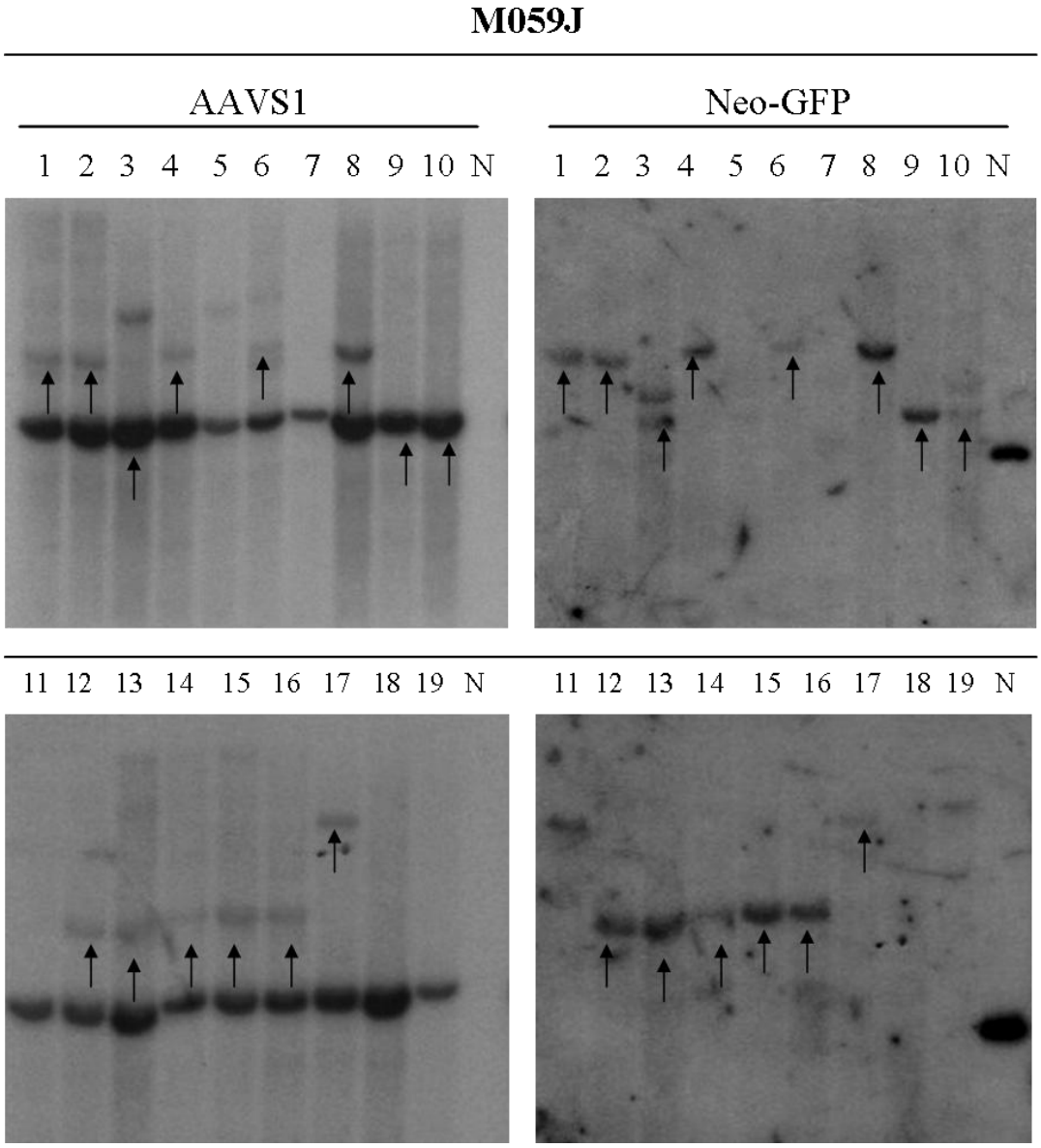



Figure 5-4: Southern hybridization of M059J clones infected with P5UF11 and pSVAV2 (50:1 ratio, 10^6 vp/cell). The left blots show signals for AAVS1 sequences and the right blots show signals for neomycin. The arrows indicate the clones that had both AAVS1 and neomycin signals co-linked. The endogenous AAVS1 band is seen in all clones. The presence of additional AAVS1 bands is indicative of rearrangements of the site, which occurs when it targeted. In some cases, the endogenous AAVS1 is co-linked with AAVS1, indicative of deletions occurring before integration. If integration occurs with minimal deletions, a higher AAVS1 band is detected. The samples are numbered. N represents a positive control plasmid which was used to determine specificity of the neomycin probe.

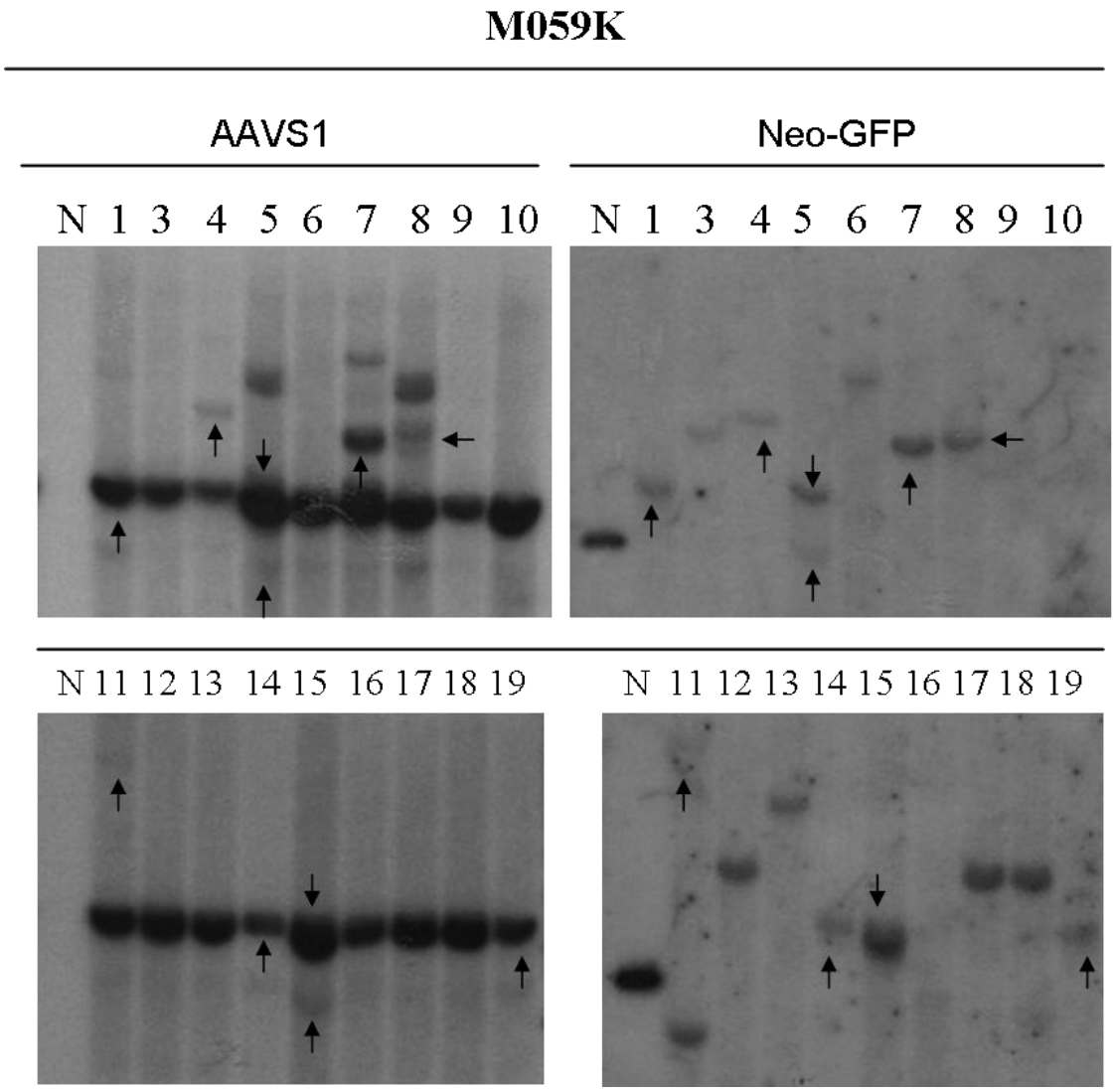



Figure 5-5. Southern hybridization of M059K clones infected with P5UF11 and pSVAV2 (50:1 ratio, 10^6 vp/cell). The left blots show signals for AAVS1 sequences and the right blots show signals for neomycin. The arrows indicate the clones that had both AAVS1 and neomycin signals co-linked. The endogenous AAVS1 band is seen in all clones. The presence of additional AAVS1 bands is indicative of rearrangements of the site, which occurs when it targeted. In some cases, the endogenous AAVS1 is co-linked with AAVS1, indicative of deletions occurring before integration. If integration occurs with minimal deletions, a higher AAVS1 band is detected. The samples are numbered. N represents a positive control plasmid which was used to determine specificity of the neomycin probe.

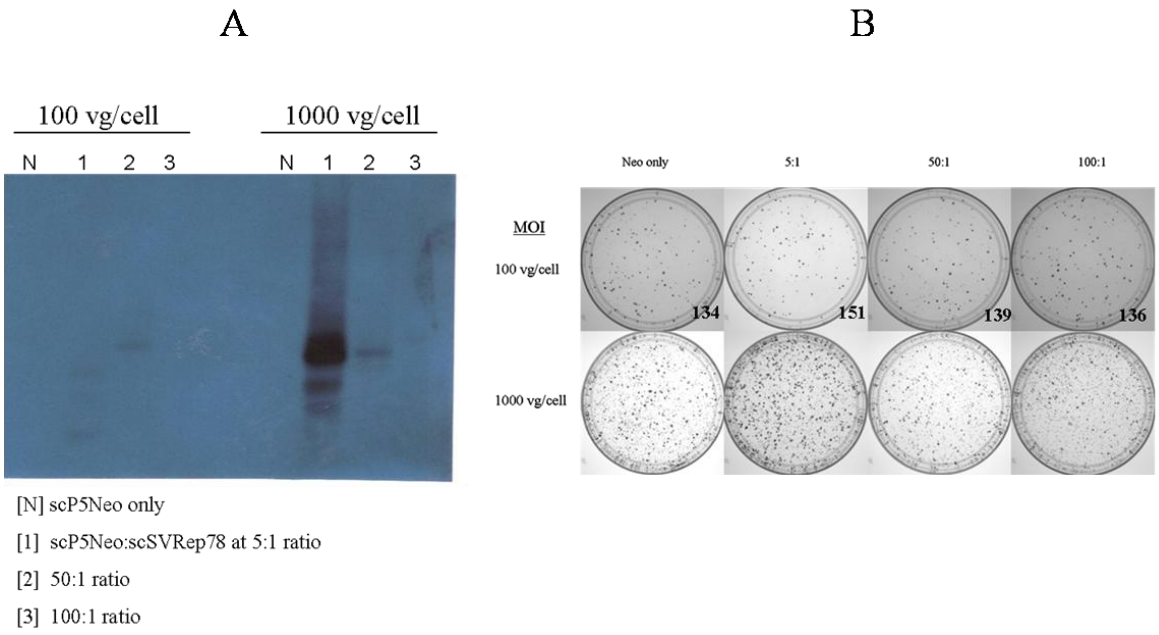



Figure 5-6: Self-Complementary AAV vectors integrate site-specifically in HeLa cells. A) Junction PCR assay on HeLa cells infected with dsP5AAVNeoR only or with dsAAV-SVRep78 at different ratios and dose (vg/cell: viral genomes per cell) analyzed one week post infection. B) Colony forming assay. 10,000 infected HeLa cells were seed into 100 mM dishes and selected with G418 for 2 weeks. Post selection the resistant colonies were stained and counted. The counts are shown only for the lower infected dose (100 vg/cell). Visually more colonies are seen at the 5:1 ratio at both doses.

In addition, a colony forming assay was used to see if Rep co-infection expression increases the number of drug resistant colonies, suggestive of targeted integration (Figure 5-6 B). In agreement, a ratio of 5:1 provided both an increase in colony formation and a strong detectable junction product. Moreover, increasing the dose of scAAV infection augmented both colony formation and junction formation. It is important to note that the increase in colony formation was not significant, suggesting that Rep does not increase the number of colony forming units. Overall, the results indicated that scAAV vectors are capable of integrating in HeLa cells and that the intrinsic integration machinery used to process ssAAV, can also process scAAV vector for integration. The next objective was to quantitate the efficiency of scAAV integration in the M059J and M059K cells.

**Self-Complementary Vectors Display Equal Integration in the Presence or Absence of DNAPKcs**

Using the optimal 5:1 ratio, M059J and M059K cells were infected with the scAAV vectors (dsP5Neo and dsSVRep78) at a total dose of 10^4 vg/cell to determine if DNAPKcs has a role in processing scAAV vectors for integration. Infected cells were selected for 2 weeks with G418. Nineteen resistant colonies were expanded in the presence of continuous selection and used for genomic DNA isolation and southern hybridizations. Interestingly, more random integration took place with scAAV vectors than expected in both cell lines; the presence or absence of DNAPkcs did not affect the frequency of site-specific integration (Figure 5-7 and Figure 5-8). Roughly 25% of clones, from both J and K cells, had scAAV integrated into AAVS1, with the other 75% randomly integrated (Table 5-2). Again, it is important to note that 10 M059K and 7 M059J clones do not show detectable neomycin. Since these clones were grown under continuous selection, these clones are presumably randomly integrated. Regardless, only 10-15% of scAAV vectors integrate site-specifically in the presence or absence of DNAPKcs.

Table 5-2. Summary of M059K and M059J self-complementary infected clones analyzed by southern hybridizations.

| | M059K Clones | M059J Clones |
|---|---|---|
| Number of clones disrupted for AAVS1: | 4 | 5 |
| Number of clone with detectable Neo: | 9 | 12 |
| Percentage of disrupted clones with Neo co-linked | 2/4 (50%) | 3/5 (60%) |
| AAVS1 specificity: | 2/9 (22%) | 3/12 (25%) |
| Random: | 7/9 (77%) | 9/12 (75%) |

**AAV-AAVS1 Junction Formation is Not Inhibited in the Absence of Ligase I and Ligase IV**

An important aspect of integration is the stable joining of the exogenous DNA to the cellular DNA. These linkages are performed by cellular repair proteins called ligases. Two such

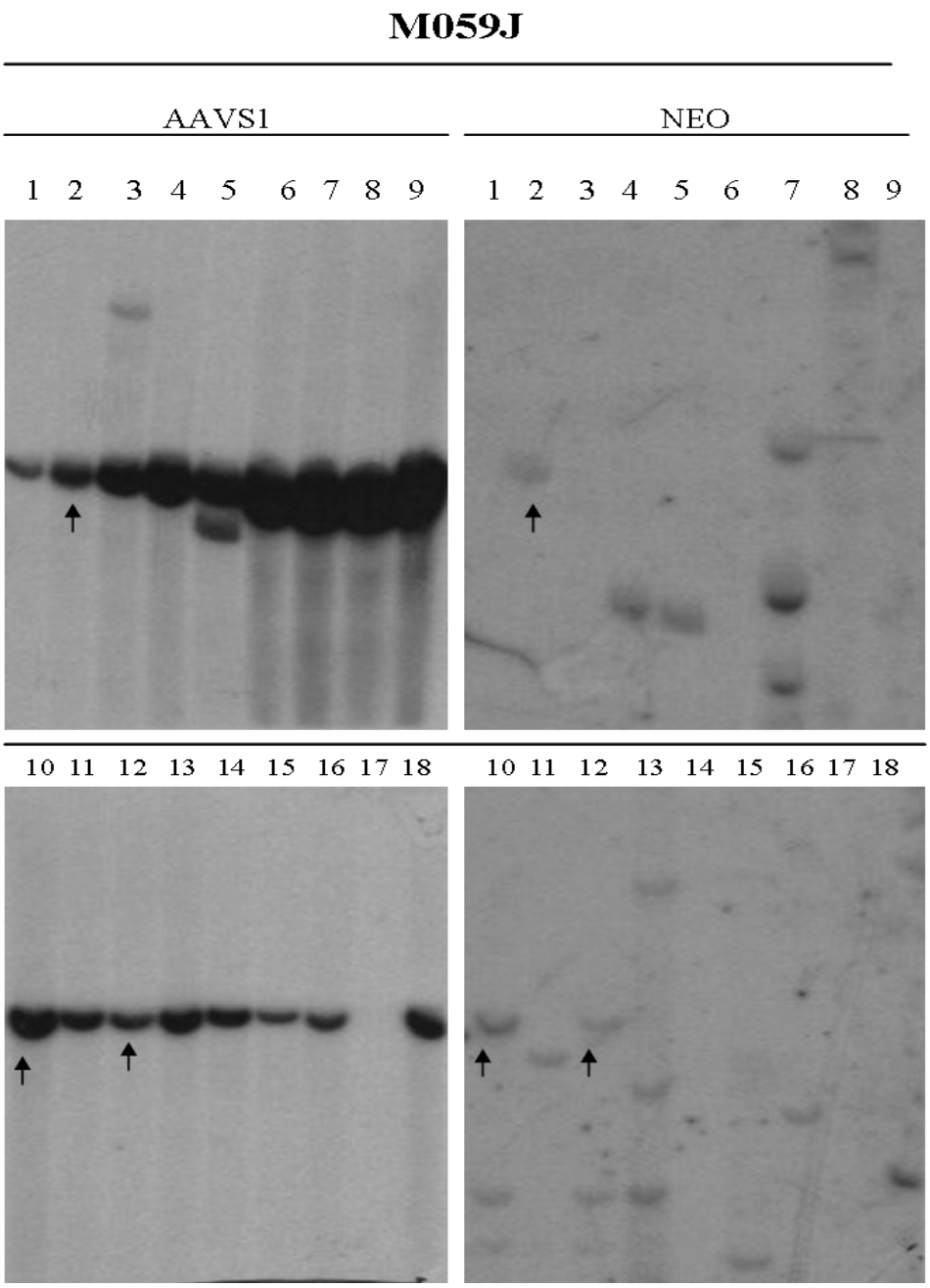



Figure 5-7. Southern hybridization of M059J clones infected with dsP5AAVNeoR and dsAAV-SVRep78 (5:1 ratio, 10^4 vg/cell). The left blots show signals for AAVS1 sequences and the right blots show signals for neomycin. The arrows indicate the clones that had both AAVS1 and neomycin signals co-linked.

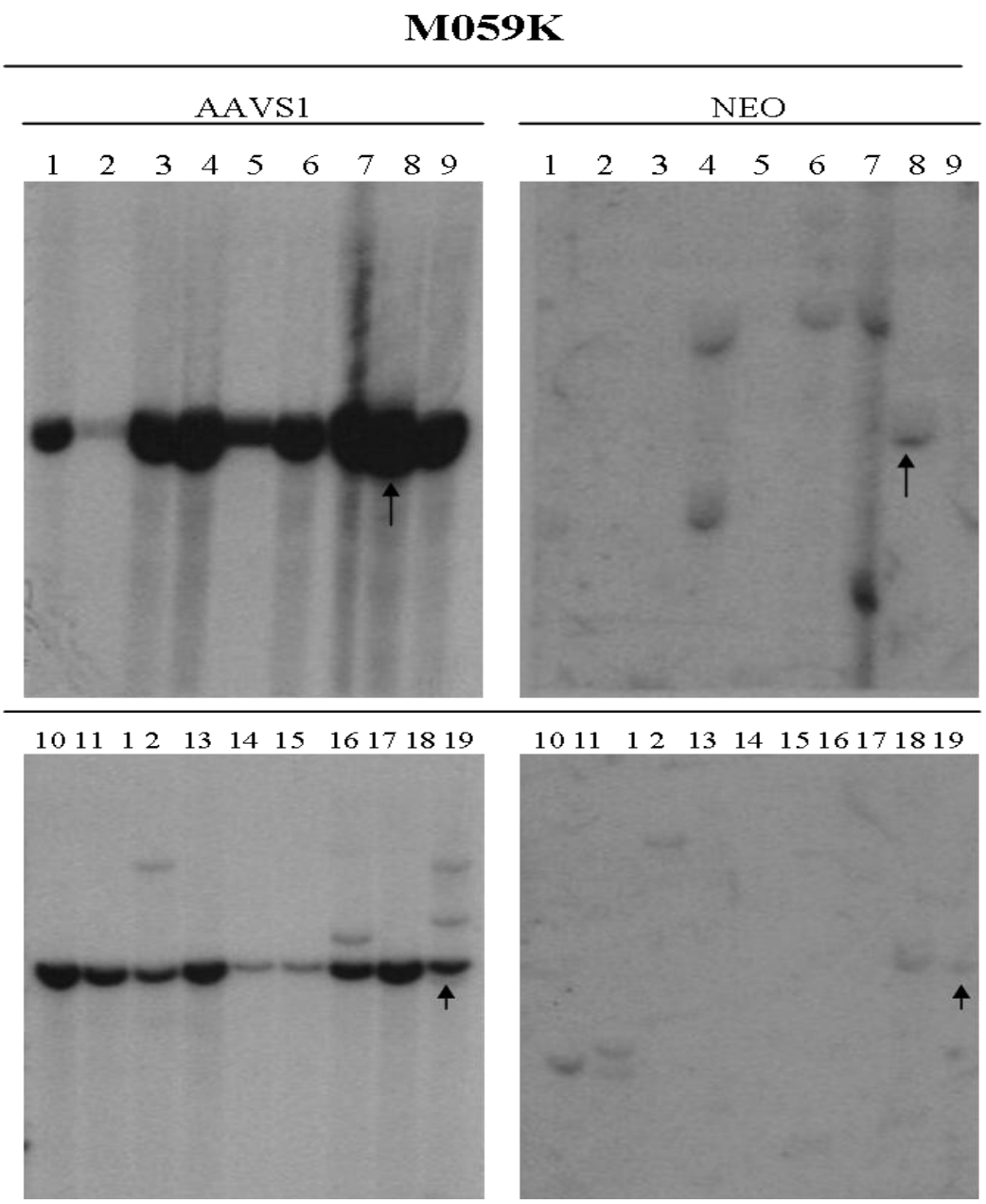



Figure 5-8. Southern hybridization of M059K clones infected with dsP5AAVNeoR and dsAAV-SVRep78 (5:1 ratio, 10^4 vg/cell). The left blots show signals for AAVS1 sequences and the right blots show signals for neomycin. The arrows indicate the clones that had both AAVS1 and Neomycin signals co-linked. band is detected. The samples are numbered. N represents a positive control plasmid which was used to determine specificity of the neomycin probe.

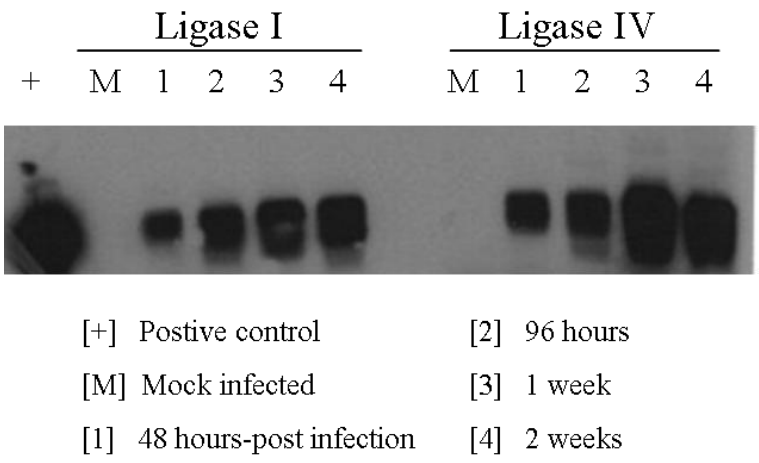



Figure 5-9. Time course of junction formation in Ligase I and Ligase IV. Both cells were infected with wild type AAV 2 at 10^5 vp/cell. Genomic DNA was isolated at the indicated time points for junction assay

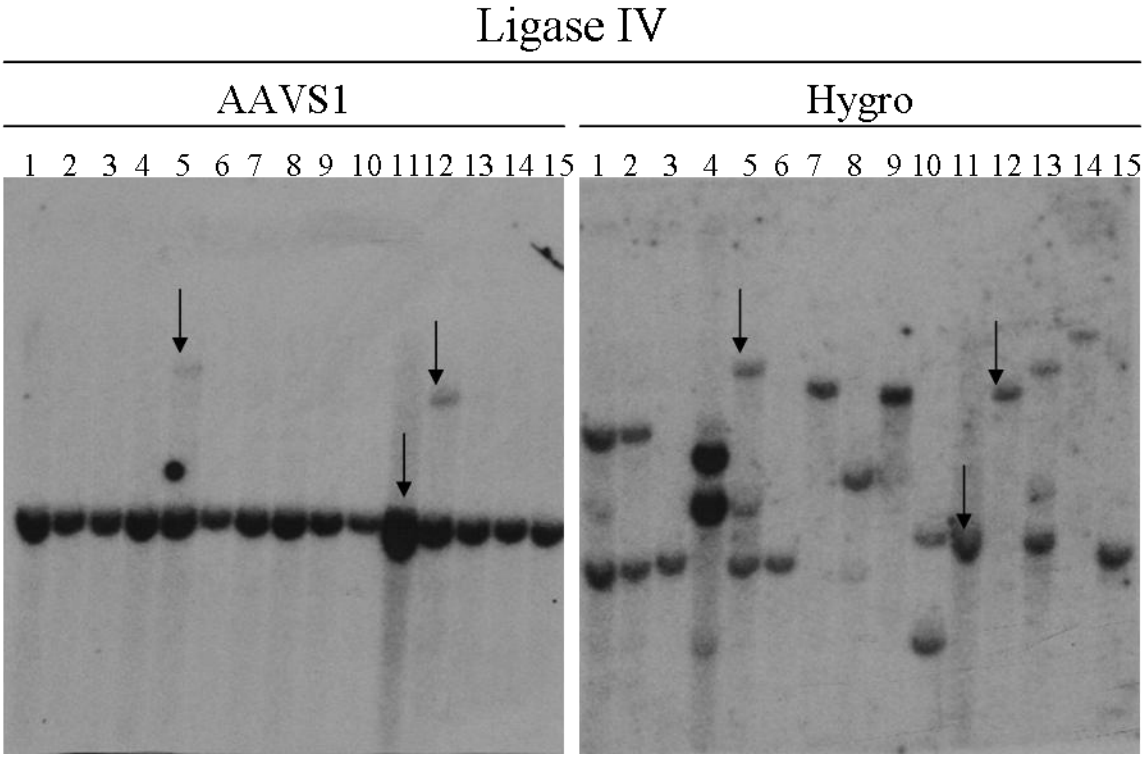



Figure 5-10. Southern hybridization of Ligase IV clones infected with P5PGKHygroGFP and pSVAV2 (50:1 ratio, 10^5 vg/cell). The left blots show signals for AAVS1 sequences and the right blots show signals for hygromycin. The arrows indicate the clones that have both AAVS1 and hygromycin signals co-linked.

ligases have been investigated for an effect on AAV-AAVS1 junction formation. ligase I and ligase IV are involved in repair by single-stranded break repair and non-homologous end joining, respectively (see reviews: Caldecott et al., 2008 and Lieber et al., 2003).

The respective ligase cells were infected with wild type AAV2 at 10^5 vp/cell. Junction formation was assessed from genomic DNA isolated form these cells 48 hrs, 96 hrs, 1 week and 2 weeks post infection. The results (Figure 5-9) clearly demonstrated that these proteins are not required for the final step of AAV integration – the ligation of the AAV DNA to the AAVS1 cellular sequence (Figure 5-9). Interestingly, southern hybridization on Ligase IV clones demonstrated that only 20% of the clones display site-specific integration (Figure 5-10). This suggests that in the absence of ligase IV the frequency of site-specific integration is greatly reduced supporting a role for ligase IV in the mechanism of AAV site-specific integration.

**Discussion And Limitation Of Study**

My data seems to suggest a role for DNAPKcs in the stability of specifically integrated single-stranded AAV vectors. First of all, a higher fraction M059K clones were not associated with specific integration as compared to M059J clones, suggesting that specific integration is less efficient in the presence of DNAPKcs. Secondly, DNAPKcs is known to be associated with Artemis, a cellular endonuclease, and it has been experimentally demonstrated that the DNAPKcs-Artemis complex can cleave hairpin loops, flaps, and gaps (Ma et al., 2005). Therefore, it is likely that either during integration and/or after integration, the AAV ITRs are cleaved by DNAPKcs –Artemis complex resulting in local chromosomal changes at the site of integration that may lead to integrant instability. This form of instability could explanation the observation by Cheung et al (1980) that more free AAV genomes (not integrated) were detected at late passage of an integrated clone (118 passages) compared to early passages (9 passages). An alternative possibility is that the integrated vector is spontaneously excised by intramolecular

recombination of the vector ITR ends. Last of all, DNAPKcs processing of input AAV genome (i.e. concatemer formation and/or circularization) could limit integration to only a small fraction of input genomes which were not processed.

*In vivo*, processing of the AAV ITR hairpins by DNA repair proteins may be tissue and cell cycle dependent. In the absence of DNAPKcs and Artemis, AAV ITR processing is impaired in muscle, heart, and kidney, but not in the liver, where DNAPKcs-independent processing took place (Inagaki et al., 2007). Furthermore, Song et al. (2004) noticed a greater percentage of transgene expression 5-weeks post-hepatectomy in SCID liver(SCID mice lack DNAPKcs expression) compared to C56/BL6 liver. Together, this suggests that processing by DNAPKcs affects maintenance of rAAV integration and expression after induction of cell proliferation. Muscle, heart, and kidney likely display different kinetics of processing and integration of rAAV genomes since they can't be induced to proliferate

The observation that self-complementary vectors showed no difference in integration in the presence or absence of DNAPKcs is quite intriguing. It partially addresses the issue of whether the substrate for site-specific integration is a single or double stranded molecule. Clearly, it suggests that hairpinned AAV vectors are much less efficient than single-stranded AAV vectors. Plasmids containing AAV sequences integrate at higher efficiencies compared to scAAV vectors, but this is likely the result of the closed, circular nature of the plasmid. Interestingly, studies have failed to detect specific integration using linear double stranded (Dyall and Berns, unpublished). Taking into account these observations, it seems that in the context of AAV infection, double stranded forms of AAV are inefficient substrates for site-specific integration, though scAAV vectors are efficient substrates for random integration, the reason for which is unclear at the moment.

It was also noticed that site-specific integration was greatly reduced in the absence of ligase IV. Coupled with the observation that junction formation was not diminished in ligase IV defective cells, several conclusions can be made with regards to analysis of specific integration and the nature of the ligase defect on integration. First of all, junction PCR is not an accurate assessment of specific integration, but just a qualitative one suggestive of integration into AAVS1; therefore, southern analysis of clones was used since it provided quantitative information with regard to frequency and specificity of integration. Secondly, although no difference in junction formation was seen between ligase I and ligase IV experiments, there may be vastly different frequencies of specific integration between the two. It is expected that specific integation should occur at normal frequency (~70%) in ligase I defective cells, since ligase I is not involved in NHEJ. Lastly, the ligase IV data strongly indicates that ligase IV is necessary for AAV site-specific integration, and therefore suggests that NHEJ repair pathway is important in AAV specific integration. The observed 20% specific integration in ligase IV defective cells may be the results of the following: (i) residual activity of ligase IV due to the nature of ligase IV mutation (Girard et al., 2004), and (ii) the ability of other ligases to participate in NHEJ mediated repair in the absence of ligase IV (Lieber et al., 2008).

There seems to be a dynamic interaction between the AAV and cellular proteins which can greatly affect expression and integration. Future research should look for Rep interaction in the absence of adenovirus using a similar experimental approach as the one performed by Nash et al. (2007). The results from such a study could serve as a source for studying identified cellular proteins for effects on AAV site specific integration. Determining how the viral Rep protein, the different AAV vectors, and the cellular components interact upon infection or integration will be crucial for developing AAV vectors with the capacity to efficiently target the AAVS1 site and

minimize risks associated with random integrations. These studies have implications for the optimal use of AAV vectors in both *ex vivo* and *in vivo* gene therapy modalities.

# CHAPTER 6
# CONCLUSION

Cellular recombination proteins clearly must have a role in AAV site specific integration because AAV is largely dependent on host factors for all aspects of its life cycle. Analysis of AAV-cellular junction sequences has revealed that the integration process is error prone, strongly suggesting the involvement of non-homologous end joining proteins and/or a short list of other repair proteins in integration.

The viral requirements for targeting site-specific integration have been elucidated, with Rep being the only viral *trans* factor required. The necessary *cis* components are the ITRs and the P5 integration efficiency element, both of which contain the RBE and TRS sequences.

Data from my first research objectives have suggested that more studies are needed to understand how Rep functions to switch AAV towards latency, and to determine if Rep can function to minimize random integration. I have been able to isolate many AAV-cellular junctions, but the limitation of the approaches used demands that alternative methods be devised and exploited to help address this issue.

My studies on the role of cellular repair proteins on AAV site-specific integration have revealed some interesting features. Firstly, DNAPKcs is not required for specific integration of single-stranded Rep containing AAV vectors. Second of all, in its absence a greater frequency of site-specific integration occurred, suggesting that DNAPKcs maybe effecting instability of the provirus after integration. This is also partially supported by the observation that more AAV-AAVS1 junction events were detected in the presence of DNAPKcs early post-infection. Lastly, self-complementary AAV vectors are capable of integration regardless of DNAPKcs, but interestingly, no difference was observed, suggesting that that DNAPKcs may function differentially in processing single-stranded versus double stranded vectors. In addition, the

frequency of site-specific integration was greatly reduced in the absence of ligase IV. Overall, these observations suggest that proteins of the non-homologous end joining pathway can modulate AAV site-specific integration.

Interestingly, pull-down assays have identified numerous DNA repair proteins that interact with Rep including Ku70, Ku80, replication protein A, Rad 50, Poly-ADP-polymerase I, DNA-dependent protein kinase (DNAPK), and proliferating cell nuclear antigen (PCNA) (Nash et al., 2009). Ku70 and Ku80 are important DNA end binding proteins in NHEJ, which recruit DNAPKcs to phosphorylate other repair proteins. Recruitment and activation of Rad50 as part of the Mre11/Rad50/NBS1 (MRN) complex can then serve to trim the DNA ends. The interaction of PCNA with cellular polymerases can then fill in the gaps, which are then ligated. Interestingly, PARP1 does not belong to NHEJ repair pathway, having a main role in single-strand break repair (SSBR). Therefore, it remains to be determined if these proteins (i.e Ku70/80, PARP1, Rad50) as well as other repair proteins, which do not interact with Rep, have roles in integration by acting on the AAV hairpins or the target chromosome 19 site.

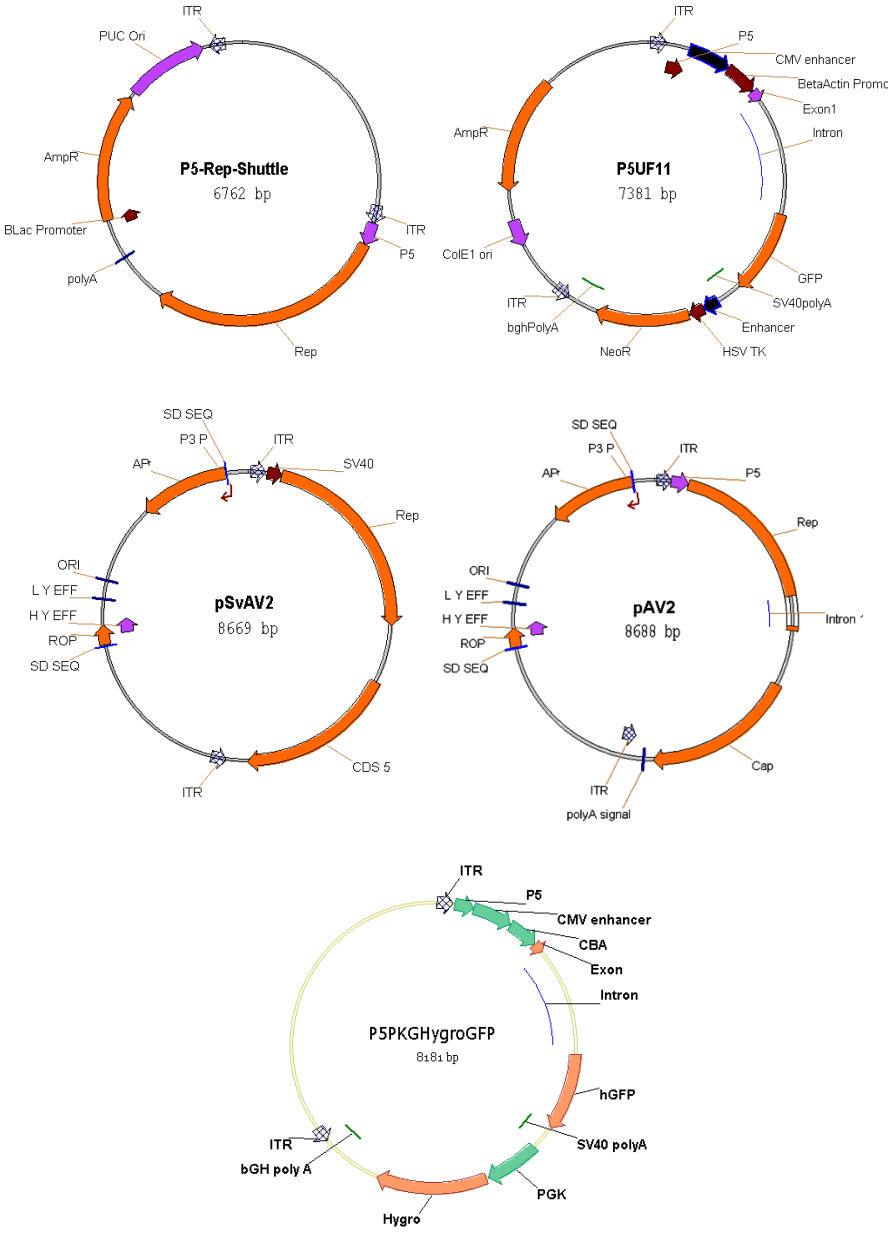
Whether by missing DNAPKcs means that other repairs pathway can now be engaged in repair mediated specific integration of AAV, remains to be studied. Thus, it is conceivable that the different cellular repair pathways are actively competing with each other for binding and processing of the AAV genome. Overall, the emerging picture is that DNA repair proteins may be shunting the AAV genome towards alternative pathways which link AAV expression and integration as competing aspects in the AAV life cycle. This is partially supported by the observations that the MRN complex binds to the AAV genomes and that knockdown of these proteins increase transduction (Cervelli et al., 2008; Schwartz et al., 2007). What could mediate the switch from expression to integration is unknown. Fragkos et al., have demonstrated that

having the P5 sequence in a recombinant vector was able to trigger DNA damage signaling (Fragkos et al., 2008), suggesting that the P5 sequence signals DNA damage response, thus recruiting a host of cellular repair proteins.

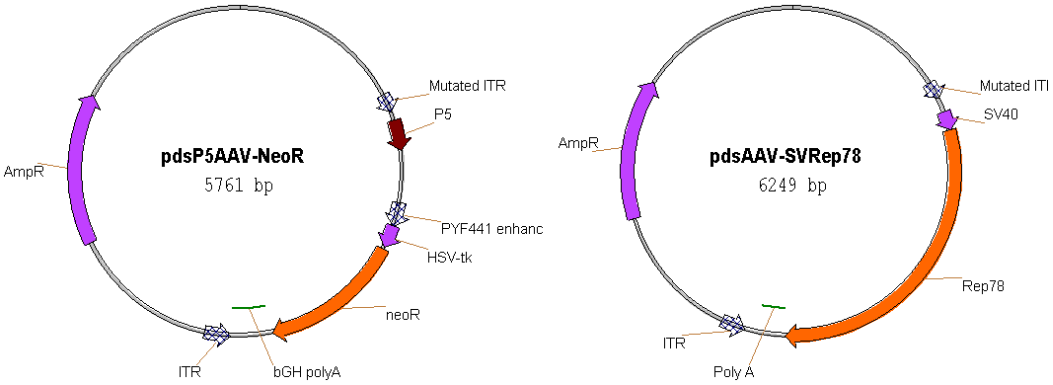
By systematically studying different cellular repair proteins for an effect on specific integration future work should be able to elucidate the mechanism AAV site-specific integration or at least identify which cellular repair proteins are required. The outcome of such advancements should lead to an increase emphasis for the development of AAV as an integrating vector that is safe for future gene therapy studies.

APPENDIX
VECTOR MAPS

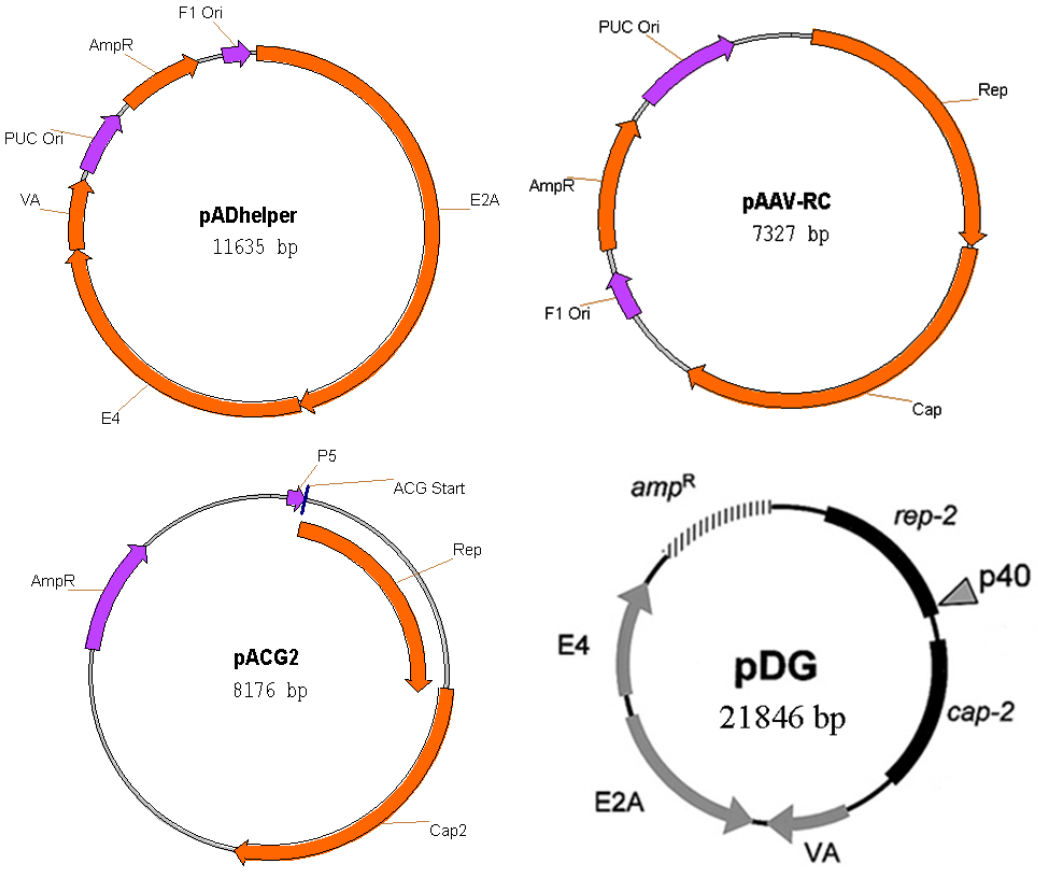
## Single Stranded Recombinant AAV Vectors



## Self-Complementary AAV Vectors




## Packaging Helper Plasmids



REFERENCES

**Afione, S.A., Conrad, C.K., Kearns, W.G., Chunduru, S., Adams, R., Reynolds, T.C., Guggino, W.B., Cutting, G.R., Carter, B.J., and Flotte, T.R.** (1996). In vivo-model of adeno-associated virus vector persistence and rescue. *J. Virol*. **70**, 3235-3241.

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

**Akache, B., Grimm, D., Shen, X., Fuess, S., Yant, S.R., Glazer, D.S., Park, J., and Kay, M.A.** (2007). A two-hybrid screen identifies cathepsins B and L as uncoating factors for adeno-associated virus 2 and 8. *Mol. Ther.* **15**, 330-339.

**Amiss, T.J., McCarty, D.M., Skulimowski, A., and Samulski, R.J.** (2003). Identification and characterization of an adeno-associated virus integration site in CV-1 cells from African green monkey. *J. Virol.* **77**, 1904-1915.

**Apparailly, F., Khoury, M., Vervoodeldonk, M.J.B., Adriaansen, J., Gicquel, E., Perez, N., Riviere, C., Louis-Plence, P., Noel, D., Danos, O., Douar, A.M., Tak, P.P., and Jorgensen, C.** (2005). Adeno-associated virus pseudotype 5 vector improves gene transfer in arthritic joints. *Hum. Gene. Ther*. **16**, 426-434.

**Ashktorab, H., and Srivastava, A.** (1989). Identification of a nuclear proteins that specifically interact with the adeno-associated virus type 2 inverted terminal repeat hairpin DNA. *J. Virol*. **63**, 3034-3039.

**Asokan, A., Hamra, J.B., Govindasamy, L., Agbandje-McKenna, M., and Samulski, R.J.** (2006). Adeno-associated virus type 2 contains an integrin . α5β1 binding domain essential for viral cell entry. *J. Virol.* **80,** 8961-8969.

**Atchison, R.W., Casto, B.C., and Hammon, W.M.** (1965). Adenovirus-associated defective virus particles. *Science*. **147**, 754-756.

**Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., Petersen-Jones, S., Bhattacharya, S.S., Thrasher, A.J., Fitzke, F.W., Carter, B.J., Rubin, G.S., Moore, A.T., and Ali, R.R.** (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N. Engl. J. Med*. **358**, 2231-2239.

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

**Bartlett, J.S., Wilcher, R., and Samulski, R.J.** (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J. Virol*. **74**, 2771-2785.

**Berns, K.I., Pinkerton, T.C., Thomas, G.F., and Hoggan, M.D.** (1975). Detection of adeno-associated virus (AAV)-specific nucleotide sequences in DNA isolated from latently infected Detroit 6 cells. *Virology*. **68**, 556-560.

**Blankinship, M., Gregorevic, P., Allen, J.M., Harper, S.Q., Harper, H., Halbert, C.L.,Miller, D.A., and Chamberlain, J.S.** (2004). Efficient transduction of skeletal muscle using vectors based on adeno-associated virus serotype 6. *Mol. Ther*. **10**, 671-678.

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

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

**Booth, M.J., Mistry, A., Li, X., Thrasher, A., and Coffin, R.S.** (2004). Transfection-free and scalable recombinant AAV vector production using HSV/AAV hybrids. *Gene. Ther*. **11**, 829-837.

**Brister, J.R., and Muzyczka, N.** (2000). Mechanism of rep-mediated adeno-associated virus origin nicking. *J. Virol*. **74**, 7762-7771.

**Buller, R.M.L., Janik, J.E., Sebring, E.D., and Rose, J.A.** (1981). Herpes simplex virus type 1 and 2 completely help adenovirus-associated virus replication. *J. Virol.* **40**, 241-247.

**Caldecott, K.W.** (2008). Single-strand break repair and genetic disease. *Nat. Rev. Genet*. **9**, 619-631.

**Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P,, Selz, F,, Hue, C,, Certain, S,, Casanova, J.L., Bousso, P., Deist, F.L., and Fischer, A.** (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease.*Science*. **288**, 669-672.

**Cervelli, T., Palacios, J.A., Zentilin, L., Mano, M., Schwartz, R.A., Weitzman, M.D., and Giacca, M.** (2008). Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. *J. Cell Sci.* **121(Pt3)**, 349-357.

C**hao, H., Liu, Y., Rabinowitz, J.,Li, C.,Samulski, R.J., and Walsh, C.E.** (2000). Several log increases in therapeutic transgene delivery by distinct adeno-associated viral serotype vectors. *Mol. Ther.* **2**, 619-623.

**Chen, S., Kapturczak, M., Loiler, S.A., Zolotukhin, S., Glushakova, O.Y., Madsen, K.M., Samulski, R.J., Hauswirth, W.W., Campbell-Thompson, M., Berns, K.I., Flotte, T.R., Atkinson, M.A., Tisher, C.C., and Agarwal, A.** (2005). Efficient transduction of vascular endothelial cells with recombinant adeno-associated virus serotype 1 and 5 vectors. *Hum. Gen. Ther*. **16,** 235-247.

**Chejanovsky, N., and Carter, B.J.** (1989). Mutagenesis of an AUG codon in adeno-associated virus rep gene: effects on viral replication. *Virology*. **173**, 120-128.

**Cheung, A.K.M., Hoggan, M.D., Hauswirth, W.W., and Berns, K.I.** (1980). Integration of the adeno-associated virus genome into cellular DNA latently infected human Detroit 6 cells. *J. Virol.* **33**, 739-748.

**Chiorini, J.A., Weitzman, M.D., Owens, R.A., Urecelay, E., Safer, B., and Kotin, R.M.** (1994). Biologically active rep proteins of adeno-associated virus type 2 produced as fusion proteins in *Escherichia coli*. *J. Virol*. **68,** 797-804.

**Choi, V.W., McCarty, D.M., and Samulski, R.J.** (2006). Host cell DNA repair pathways in adeno-associated viral genome processing. *J. Virol*. **80**, 10346-10356.

**Choi , V.W., Samuski, R.J., and McCarty, D.M.** (2005). Effects of adeno-associated virus DNA hairpin structure on recombination. *J. Virol.* **79**, 6801-6807.

**Costello, E., Saudan, P., Winocour, E., Pizer, L., and Beard, P.** (1997). High mobility group chromosomal protein 1 binds to the adeno-associated virus replication protein (Rep) and promotes rep-mediated site-specific cleavage of DNA, ATPase activity and transcriptional repression. *EMBO J.* **16**, 5943-5954.

**Davidoff, A.M., Gray, J.T., Ng, C.Y., Zhang, Y., Zhou, J., Spence, Y., Bakar, Y., and Nathwani, A.C**.(2004). Comparison of the ability of adeno-associated viral vectors pseudotyped with serotypes 2, 5, and 8 capsid proteins to mediate efficient transduction of the liver in murine and nonhuman primate models. *Mol. Ther*. **11**, 875-888.

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

**Drew, H.R., Lockett, L.J., and Both, G.W**. (2007). Increased complexity of wild type adeno-associated virus-chromosomal junctions as determined by analysis of unselected cellular genomes. *J. Gen. Virol*. **88**, 1722-1732.

**Duan, D., Li, Q., Kao, A.W., Yue, Y., Pessin, J.E., and Engelhardt, J.F.** (1999). Dynamin is required for recombinant adeno-associated virus type 2 infection. *J. Virol.* **73**, 10371-10376.

**Duan, D., Sharma, P., Yang, J., Yue, Y., Dudus, L., Zhang, Y., Fisher, K.J., and Engelhardt, J.F** (1998). Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J. Virol.* **72**, 8568-8577.

**Duan, D., Yue, Y., and Engelhardt, J.F.** (2003). Consequences of DNA-dependent protein kinase catalytic subunit deficiency on recombinant adeno-associated virus genome circularization and heterodimerization in muscle tissue. *J. Virol.* **77,** 4751-4759.

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

**Dutheil, N., Shi, F., Dupressoir, T., and Linden, R.M**. (2000). Adeno-associated virus site-specifically integrates into a muscle specific DNA region. *Proc. Natl. Acad. Sci. USA*. **97**, 4862-4866.

**Dutheil, N., Yoon-Roberts, M., Ward, P., Henckaerts, E., Skrabanek, L., Berns, K.I., Campagne, F., and Linden, R.M.** (2004). Characterization of the mouse adeno-associated virus AAVS1 ortholog. *J. Virol*. **78**, 8917-8921.

**Dyall, J., Szabo, P., and Berns, K.I.** (1999). Adeno-associated virus (AAV) site-specific integration: formation of AAV-AAVS1 junctions in an in vitro system. *Proc. Natl. Acad. Sci. USA*. **96**, 12849-12854.

**Feng, D., Chen, J., Yue, Y., Zhu, H., Xue, J., and Jia, W. W**. (2006). A 16 bp rep binding element is sufficient for mediating rep-dependent integration into AAVS1. *J Mol Biol*. **358**, 38-45

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

**Fragkos, M., Breuleux, M., Clement, N., and Beard, P.** (2008). Recombinant adeno-associated viral vectors are deficient in provoking a DNA damage response. *J. Virol*. **82**, 7379-7387.

**Gao, G.P., Alvira, M.R., Wang, L., Clacedo, R., Johnston, J., and Wilson, J.M.** (2002). Novel adeno-associated viruses from rhesus monkey as vectors for human gene therapy. *Proc. Natl. Acad. Sci. USA*. **99**, 11854-11859.

**Gao, G.P., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Xhou, X., and Wilson, J.M.** (2004). Clades of adeno-associated viruses are widely disseminated in human tissues. *J. Virol*. **78**, 6381-6388.

**Geoffrey, M.C., Epstein, A.L., Toublanc, E., Moullier, P., and Salvetti, A.** (2004). Herpes simplex virus type 1 ICP0 protein mediates activation of adeno-associated virus type 2 rep gene expression from a latent integrated from. *J. Virol*. **78**, 10977-10986.

**Ghosh, A., Yue, Y., and Duan, D.** (2006). Viral serotype and transgene sequence overlapping adeno-associated virus (AAV) vector-mediated gene transfer in skeletal muscle. *J. Gen. Med*. **8**, 298-305.

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

**Girard, P-M., Kysela, B., Harer, C.J., Doherty, A.J., and Jeggo, P.A.** (2004). Analysis of DNA ligase IV mutations found in LIG4 syndrome patients: the impact of two linked polymorphisms. *Hum. Mol. Genet*. **13**, 2369-2376.

**Giraud, C., Winocour, E., and Berns, K.I.** (1994). Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. *Proc. Natl. Acad. Sci. USA.* **91**, 10039-10043.

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

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

**Grimm, D., Kay, M.A., and Kleinschmidt, J.A.** (2003). Helpervirus-free, optically controllable, and two-plasmid-basedproduction of adeno-associated virus vectors of serotypes 1 to 6. *Mol Ther*. **7**, 839–850.

**Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E,, Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B.H., Wintergerst, U., Velez, M.C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F.D., Fischer, A., and Cavazzana-Calvo, M..** (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **118**, 3132-3142.

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

**Hansen, J., Qing, K., and Srivastava, A**. (2001). Infection of purified nuclei by adeno-associated virus 2. *Mol. Ther.* **4**, 289-296.

**Hauswirth, W.W., Aleman, T.S., Kaushal, S., Cideciyan, A.V., Schwartz, S.B., Wang, L., Conlon, T.J., Boye, S.L., Flotte, T.R., Byrne, B.J., and Jacobson, S.G.** (2008). Treatment of Leber Congenital Amaurosis Due to RPE65 Mutations by Ocular Subretinal Injectionof Adeno-Associated Virus Gene Vector: Short-Term Results of a Phase I Trial. *Hum. Gene. Ther.* **19**, 979-990.

**Henckaerts, E., Dutheil, N., Zeltner, N., Kattman, S., Kohlbrenner, E., Ward, P., Clement, N., Rebollo, P., Kennedy, M., Keller, G.M., and Linden, R.M.** (2009). Site-specific integration of adeno-associated virus involves partial duplication of the target locus. (in press).

**Hermonat, P.L., and Muzyczka, N.** (1984).Use of adeno-associated virus as a mammalian DNAcloning vector: transduction of neomycin resistance into mammalian tissue cultured cells. *Proc Natl Acad Sci USA*. **81**, 6466-6470.

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

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

**Hoggan, M.D., Shatkin, A.J., Blacklow, N. R., Koczot, F., and Rose, J.A.** (1968). Helper-dependent infectious deoxyribonucleic acid from adenovirus-associated virus. *J. Virol*. **2**, 850-851.

**Huser, D., Weger, S., and Heilbronn, R.** (2002). Kinetics and frequency of adeno-associated virus site-specific integration into human chromosome 19 monitored by quantitative real-time PCR. *J. Virol.* **76**, 7554-7559.

**Inagaki, K., Fuess, S., Storm, T.A., Gibson, G.A., McTiernan, C.F., Kay, M.A., and Nakai, H.** (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol. Ther*. **14**, 45-53.

**Inagaki, K., Lewis, S.M., Wu, X., Ma, C., Munroe, D.J., Fuess, S., Storm, T.A., Kay, M.A., and Nakai, H.** (2007). DNA palindromes with modest arm length of greater, similar 20 base pairs are a significant target for recombinant adeno-associated virus vector integration in the liver, muscles, and heart in mice. *J. Virol.* **81**, 11290-11303.

**Inagaki, K., Ma, C., Storm, T.A., Kay, M.A., and Nakai, H.** (2007). The role of DNA-PKcs and artemis in opening viral DNA hairpin termini in various tissues in mice. *J. Virol.* **81**, 11304-11321.

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

**Janik, J.E., Huston, M.M., and Rose, J.A.** (1981). Locations of adenovirus genes required for the replication of adenovirus-associated virus. *Proc. Natl. Acad. Sci. USA.* **78**, 1925-1929.

**Johnson, J.S., and Samulski, R.J.** (2009). Enhancement of AAV interaction by mobilizing capsids into and out of the nucleolus. *J. Virol*. 83, 2632-2644.

**Jurvansuu, J., Raj, K., Stasiak, A., and Beard, P.** (2005). Viral transport of DNA damage that mimics a stalled replication fork. *J. Virol.* **79**, 569-580.

**Kaludov, N., Brown, K.E., Walters, R.W., Zabner, J., and Chiorini, J.A.** (2001). Adeno-associated virus serotype 4 (AAV 4) and AAV5 both require sialic acid for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J. Virol.* **75**, 6884-6893.

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

**Kearns, W.G., Afione, S.A., Fulmer, S.B., Pang, M.C., Erickson, D., Egan, M., Landrum, M.J., Flotte, T.R., and Cutting, G.R.** (1996). Recombinant adeno-associated virus (AAV-CFTR) vectors do not integrate in a site-specific fashion in an immortalized epithelial cell line. *Gene Ther*. **3**, 748-755.

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

**Kotin, R.M., Linden, R.M., and Berns, K.I.** (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *Embo J*. **11**, 5071-5078.

**Kotin, R.M., Siniscalco, M., Samulski, R.J., Zhu, X., Hunter, L., Laughlin, C.A., McLaughlin, S., Muzyczka, N., Roccchi, M., and Berns, K.I.** (1990). Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. USA*. **87**, 2211-2215.

**Kyöstiö, S.R.M., Wonderling, R.S., and Owens, R.A.** (1995). Negative regulation of the adeno-associated virus (AAV) P5 promoter involves both the P5 rep binding site and the consensus ATP-binding motif of the AAV Rep68 protein. *J. Virol*. **69**, 6787-6796.

**Lackner, D.F., and Muzyczka, N.** (2002). Studies of the mechanism of transactivation of the adeno-associated virus p19 promoter by Rep protein. *J. Virol*. **76**, 8225-8235.

**Lamartina, S., Sporeno, E., Fattori, E., and Toniatti, C.** (2000). Characteristics of the adeno-associated virus preintegration site in human chromosome 19: open chromatin conformation and transcription competent environment. *J. Virol*. **74**, 7671-7677.

**Lieber, M.R.** (2008). The mechanism of human nonhomologous DNA end joining. *J. Biol. Chem.* **283**, 1-5.

**Lieber, M.R., Ma, Y., Pannicke, U., and Schwarz, K.** (2003). Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol*. **4**, 712-720.

**Li, J., Samulski, R.J., and Xiao, X.** (1997). Role for highly regulated rep gene expression in adeno-associated virus vector production. *J. Virol.* **71**, 5236-5243.

**Li, Z., Brister, J.R., Im, D.S., and Muzyczka, N**. (2003). Characterization of the adeno-associated virus rep protein complex formed on the viral origin of replication. *Virology*. **313**, 364-376.

**Linden, R.M., Winocour, E., and Berns, K.I.** (1996). The recombination signals for adeno-associated virus site-specific integration. *Proc. Natl. Acad. Sci. USA*. **93**, 7966-7972.

**Liu, X., Luo, M., Zhang,L.N., Yan, Z., Zak, R., Ding, W., Mansfield, S.G., Mitchell, L.G., and Engelhardt, J.F.** (2005). Spliceosome-mediated RNA trans-splicing with recombinant adeno-associated virus partially restores cystic fibrosis transmembrane conductance regulator function to polarized human cystic fibrosis airway epithelial cells. *Hum. Gene Ther.* **16**, 1116-1123.

**Louboutin, J.P., Wang, L., and Wilson, J.M.** (2005). Gene transfer into skeletal muscle using using novel AAV serotypes. *J. Gene Med.* **7**, 442-451.

**Lux, K., Goerlitz, N., Schlemminger, S., Perabo, L., Goldnaw, D., Endell, J., Leike, K., Kofler, D.M., Finke, S., Hallek, M., and Büning, H.** (2005). Green fluorescent protein-tagged adeno-associated virus particles allow the study of cytosolic and nuclear trafficking. *J. Virol.* **79**, 11776-11787.

**Ma, Y., Schwarz, K., and Lieber, M.R.** (2005). The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amst).* **4**, 845-851.

**Maguire, A.M., Simonelli, F., Pierce, E.A., Pugh, E.N. Jr,, Mingozzi, F., Bennicelli, J., Banfi, S., Marshall, K.A., Testa, F., Surace, E.M., Rossi, S., Lyubarsky, A., Arruda, V.R., Konkle, B., Stone, E., Sun, J., Jacobs, J., Dell'Osso, L., Hertle, R., Ma, J.X., Redmond, T.M., Zhu, X., Hauck, B., Zelenaia, O., Shindler, K.S., Maguire, M.G., Wright, J.F., Volpe, N.J., McDonnell, J.W., Auricchio, A., High, K.A., and Bennett, J.** (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med.* **358**, 2240-2248.

**Mah, C., Qing, K., Khuntirat, B., Ponnazhagan, S. Wang, X.S., Kube, D.M., Yoder, M.C., and Srivastava, A.** (1998). Adeno-associated virus type 2-mediated gene transfer: role of epidermal growth factor receptor protein tyrosine kinase in transgene expression. *J. Virol.* **72,** 9835-9843.

**Manno, C.S., Pierce, G.F., Arruda, R., Glader, B., Ragni, M., Rasko, J.J., Ozelo, M.C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Rustagi, P.K., Nakai, H., Chew, A., Leonard, D., Wright, J.F., Lessard, R.R., Sommer, J.M., Tigges, M., Sabatino, B., Luk, A., Jiang, H., Mingossi, F., Couto, L., Ertl, H.C., High, K.A., and Kay, M.A.** (2006). Successful transduction of liver in hemophilia by AAV-factor IX and limitations imposed by the host immune response. *Nat. Med.* **12**, 342-347.

**Mayor, H.D., Torikai, K., Melnick, J.L., and Mandel, M.** (1969). Plus and minus single-stranded DNA separately encapsidated in adeno-associated satellite virions. *Science.* **166**, 1280-1282.

**McCarty, D.M., Fu, H., Monahan, P.E., Toulson, C.E., Naik, P., and Samulski, R.J.** (2003). Adeno-associated virus terminal repeat (TR) mutants generate self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene. Ther*. **10**, 2112-2118.

**McCarty, D.M., Monahan, P.E., and Samulski, R.J.** (2001). Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther*. **16**, 1248-1254.

**McLaughlin, S.K., Collis, P., Hermonat, P.L., and Muzyczka, N.** (1988). Adeno-associated virus general transduction vectors: analysis of proviral structures. *J. Virol*. **62**, 1963-1973.

**Mehrle, S., Rohde, V., and Schlehofer, J.R.** (2004). Evidence of chromosomal integration of AAV DNA in human testis tissue. *Virus Genes.* **28**, 61-69.

**Melnick, J.L., Mayor, H.D., Smith, K.O., and Rapp, F.** (1965). Association of 20-millimicron particles with adenovirus. *J. Bacteriol*. **90**, 271-274.

**Meneses, P., Berns, K.I., and Winocour, E.** (2000). DNA sequence motifs which direct adeno-associated virus site-specific integration in a model system. *J. Virol.* **74**, 6213-6216.

**Merten, O-W., Geny-Fiamma, C., and Douar, A.M.** (2005). Current issues in adeno-associated viral vector production. *Gene. Ther*. **12**, S51-S61.

**Miller, D.G., Trobridge, G.D., Petek, L.M., Jacobs, M.A., Kaul, R., and Russell, D.W.** (2005). Large-scale analysis of adeno-associated virus vector integration sites in normal human cells. *J. Virol.* **79**, 11434-11442.

**Mingozzi, F., Maus, M.V., Hui, D.J., Sabatino, D.E., Murphy, S.L., Rasko, J.E., Ragni, M.V., Manno, C.S., Sommerm J., Jiang, H., Pierce, G.F., Ertl, H.C., and High, K.A.** (2007). CD8 (+) T-cell responses to adeno-associated virus capsids in humans. *Nat. Med*. **13**, 419-422.

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

**Murphy, M., Gomos-Klein, J., Stankic, M., and Falck-Pedersen, E.** (2007). Adeno-associated virus type 2 p5 promoter: a rep-regulated DNA switch element functioning in transcription, replication, and site-specific integration. *J. Virol.* **81**, 3721-3730.

**Nakai, H., Fuess, S., Storm, T.A., Meuse, L.A., and Kay, M.A.** (2003). Free DNA ends are essential for concatemerization of synthetic double-stranded adeno-associated virus vector genomes transfected into mouse hepatocytes in vivo. *Mol. Ther.* **7**, 112-121.

**Nakai, H., Fuess, S., Storm, T.A, Muramatsu, S., Nara, Y., and Kay, M.A.** (2005). Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors I mice. *J. Virol*. **79**, 214-224.

**Nakai, H., Yuichi, I., Kay, M.A., and Couto, L.B**. (1999). Isolation of recombinant adeno-associated virus vector cellular DNA junctions from mouse liver. *J. Virol*. **73**, 5438-5447.

**Nash, K., Chen, W., McDonald, W.F., Zhou, X., and Muzyczka, N.** (2007). Purification of host cells enzymes involved in adeno-associated virus DNA replication. *J. Virol*. **81**, 5777-5787.

**Nash, K., Chen, W., and Muzyczka, N.** (2008). Complete in vitro reconstitution of adeno-associated virus DNA replication requires mini-chromosome maintenance complex proteins. *J. Virol*. **82,** 1458-1464.

**Nash, K., Chen, W., Salganik, M., and Muzyczka, N**. (2009). Identification of cellular proteins that interact with the adeno-associated virus rep protein. *J. Virol*. **83**, 454-469.

**Ogata, T., Kozuka, T., and Kanda, T.** (2003). Identification of an insulator in AAVS1, a preferred region for integration of adeno-associated virus DNA. *J. Virol.* **77**, 9000-9007.

**Opie, S.R., Warrington Jr., K.H., Agbandje-McKenna, M., Zolotukhin, S., and Muzyczka, N.** (2003). Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. *J. Virol*. **77**, 6995-7006.

**Pacak, C.A., Mah, C.S., Thattaliyath, B.D., Conlon, T.J., Lewis, M.A., Cloutier, D.E., Zoltukhin, I., Tarantal, A.F., and Byrne, B.J.** (2006). Recombinant adeno-associated virus serotype 9 leads to preferential cardiac transduction in vivo. *Circ. Res.* **99**, e3-e9.

**Pasquale, G.D., Davidson, B.L., Stein, C.S., Martins, I., Scudiero, D., Monks, A., and Chiorini, J.A.** (2003). Identification of PDGFR as a receptor for AAV-5 transduction. *Nat. Med.* **9**, 1306-1312.

**Philpott, N.J., Gomos, J., Berns, K.I., and Falck-Pedersen, E.** (2002). A p5 integration efficiency element mediates rep-dependent integration into AAVS1 at chromosome 19. *Proc. Natl. Acad. Sci. USA*. **99**, 12381-12385.

**Philpott, N.J., Gomos, J., and Falck-Pedersen, E.** (2004). Transgene expression after rep-mediated site-specific integration into chromosome 19. *Hum. Gene. Ther*. **15**, 47-61.

**Ponnazhagan, S., Erikson, D., Kearns, W.G., Zhou, S.Z., Nahreini, P., Wang, X.S., and Srivastava, A.** (1997). Lack of site-specific integration of the recombinant adeno-associated virus 2 genomes in human cells. *Hum. Gene. Ther.* **8**, 275-284.

**Prasad, K.M.R., and Trempe, J.P.** (1995). The adeno-associated virus Rep78 protein is covalently linked to viral DNA in preformed virions. *Virology*. **214**, 360-370.

**Qing, K., Hansen, J., Weigel-Kelley, K.A., Tan, M., Zhou, S., and Srivastava, A.** (2001). Adeno-associated virus type 2 mediated gene transfer: role of cellular FKBP52 protein in transgene expression. *J. Virol*. **75**, 8968-8976.

**Qing, K., Li, W., Zhong, L., Tan, M., Hansen, J., Weigel-Kelley, K.A., Chen, L., Yoder, M.C., and Srivastava, A**. (2003). Adeno-associated virus type 2-mediated gene transfer: role of cellular T-cell protein tyrosine phosphatase in transgene expression in established cell lines in vitro and and transgenic mice in vivo. *J. Virol.* **77**, 2741-2746.

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

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

**Raper, S.E., Chirmule, N., Lee, F.S., Wivel, N.A., Bagg, A., Gao, G.P., Wilson, J.M., and Batshaw, M.L.** (2003). Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol Genet Metab*. **80**, 148-158.

**Raper, S.E., Yudkoff, M., Chirmule, N., Gao, G.P., Nunes, F., Haskal, Z.J., Furth, E.E., Propert, K.J., Robinson, M.B., Magosin, S., Simoes, H., Speicher, L., Hughes, J., Tazelaar, J., Wivel, N.A., Wilson, J.M., and Batshaw, M.L.** (2002). A pilot study of in vivo liver-directed gene transfer with an adenoviral vector in partial ornithine transcarbamylase deficiency. *Hum. Gene Ther.* **13**, 163-175.

**Reich, S.J., Auricchio, A., Hildinger, M., Glover, E., Maguire, A.M., Wilson, J.M., and Bennett, J.** (2003). Efficient trans-splicing in the retina expands the utility of adeno-associated virus as a vector for gene therapy. *Hum. Gene Ther*. **14**, 37-44.

**Ren, C., White, A.F., and Ponnazhagan, S.** (2007). Notch1 augments intracellular trafficking of adeno-associated virus type 2. *J. Virol.* **81**, 2069-2073.

**Rose, J.A, Berns, K.I., and Hoggan, M.D., and Koczot, F.J.** (1969). Evidence for single-stranded adenovirus-associated virus genome: formation of DNA density-hybrid on release of viral DNA. *Proc. Natl. Acad. Sci. USA.* **64**, 863-869.

**Salvetti, A., Oreve, S., Chadeuf, G., Favre, D., Cherel, Y., Champion-Arnaud, P., David-Ameline, J., and Moullier, P.** (1998). Factors influencing recombinant adeno-associated virus production. *Hum Gene Ther.* **9**, 695–706.

**Samulski, R.J., Zhu, X., Xiao, X., Brook, J.D., Housman, D.E., Epstein, N., and Hunter, L.A**. (1991). Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *Embo J.* **10**, 3941-3950.

**Sanlioglu, S., Benson, P.K., and Engelhardt, J.F. (**2000). Loss of ATM function enhances recombinant adeno-associated virus transduction and integration through pathways similar to UV irradiation. *Virology*. **268**, 68-78.

**Sanlioglu, S., Benson, P.K., Yang, J., Atkinson, E.M., Reynolds, T., and Engelhardt, J.F.** (2000). Endocytosis and nuclear trafficking of adeno-associated virus type 2 are controlled by Rac1 and phosphatidylinositol 3 kinase activation. *J. Virol.* **74**, 9184-9196.

**Schnepp, B.C., Clark, K.R., Klemanski, D.L., Pacak, C.A., and Johnson, P.R.** (2003). Genetic fate of recombinant adeno-associated virus vector genomes in muscle. *J. Virol*. **77**, 3495-3504.

**Schnepp, B.C., Jensen, R.L., Chen, C.L., Johnson, P.R., and Clark, K.R.** (2005). Characterization of adeno-associated virus genomes isolated from human tissues. *J. Virol*. **79**, 14793-14803.

**Schnepp, B.C., Jensen, R.L., Clark, K.R., and Johnson, P.R.** (2009). Infectious molecular clones of adeno-associated virus isolated directly from human tissues. *J. Virol*. **83**, 1456-1464.

**Schwartz, R.A., Palacios, J.A., Cassell, G.D., Adam, S., Giacca, M., and Weitzman, M.D.** (2007). The Mre11/Rad50/Nbs1 complex limits adeno-associated virus transduction and replication. *J. Virol*. **81**, 12936-12945.

**Seisenberger, G., Ried, M.U., Endre, T., Büning, H., Hallek, M., and Bräuchle, C.** (2001). Real time single-molecule imaging of the infectious pathway of an adeno-associated virus. *Science*. **294**, 1929-1932.

**Snyder, R.O., Im, D.S., Ni, T., Xiao, X., Samulski, R.J., and Muzyczka, N**. (1993). Features of the adeno-associated virus origin involved in substrate recognition by the viral rep protein. *J. Virol.* **67**, 6096-6104.

**Sollerbrant, K., Elmen, J., Wahlestedt, C., Acker, J. Leblois-Prehaud, H., Latta-Mahieu, M., Yeh, P., and Perricaudet, M.** (2001). A novel method using baculovirus-mediated gene transfer for production of recombinant adeno-associated virus vectors. *J. Gen. Virol.* **82**, 2051-2060.

**Song, S., Lu, Y., Choi, Y.K., Han, Y., Tang, Q., Zhao, G., Berns, K.I., and Flotte, T.R.** (2004). DNA-dependent PK inhibits adeno-associated virus DNA integration. *Proc. Nat. Acad. Sci. USA.* **101**, 2112-2116.

**Srivastava, A., Lusby, E.W., and Berns, K.I**. (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J. Virol*. **45**, 555-564.

**Summerford, C., Bartlett, J.S., and Samulski, R.J.** (1999). αVβ5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat. Med*. **5**, 78-82.

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

**Surosky, R.T., Urabe, M., Godwin, S.G., McQuiston, S.A., Kurtzman, G.J., Ozawa, K., and Natsoulis, G.** (1997). Adeno-associated virus rep proteins target DNA sequences to a unique locus in the human genome. *J. Virol.* **71**, 7951-7959.

**Tamayose, K., Hirai, Y., and Shimada, T.** (1996). A new strategy for large-scale preparation of high-titer recombinant adeno-associated virus vectors by using packaging cell lines and sulfonated cellulose column chromatography. *Hum. Gen. Ther.* **7**, 504-513.

**Tan, I., Ng, C.H., Lim, L., and Leung, T.** (2001). Phosphorylation of a novel myosin binding subunit of protein phosphatase 1 reveals a conserved mechanism in the regulation of actin cytoskeleton. *J. Biol. Chem.* **276**, 21209-2126.

**Urcelay, E., Ward, P., Wiener,S.M., Safer, B., and Kotin, R.M.** (1995). Asymmetric replication in vitro from a human sequence element is dependent on adeno-associated virus rep protein. *J. Virol*. **69**, 2038-2046.

**Vincent-Lacaze, N., Snyder, R.O., Gluzman, R., Bohl, D., Lagarde, C., and Danos, O.** (1999). Structure of adeno-associated virus vector DNA following transduction of the skeletal muscle. *J. Virol.* **73**, 1949-1955.

**Wang, C., Wang, C.M., Clark, K.R., and Sferra, T.J.** (2003). Recombinant AAV serotype 1 transduction efficiency and tropism in the murine brain. *Gene. Ther.* **10**, 1528-1534.

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

**Wang, X.S, Qing, K., Ponnazhagan, S., and Srivastava, A**. (1997). Adeno-associated virus type 2 DNA replication in vivo: mutation analyses of the D sequence in viral inverted terminal repeats. *J. Virol.* **71**, 3077-3082.

**Wang, Z., Zhu, T., Qiao, C., Zhou, L., Wang, B., Zhang, J., Chen, C., Li, J., and Xiao, X.** (2005). Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. *Nat. Biotechnol.* **23**, 321-328.

**Ward, P., Falkenberg, M., Elias, P., Weitzman, M.D., and Linden, R.M.** (2001). Rep-dependent initiation of adeno-associated virus type 2 DNA replication by a herpes simplex virus type 1 replication complex in a reconstituted system. *J. Virol.* **75**, 10250-10258.

**Weger, S., Wendland, M., Kleinschmidt, J.A, and Heilbronn, R.** (1999). The adeno-associated virus type 2 regulatory proteins Rep78 and Rep68 interact with the transcriptional coactivator PC4. *J. Virol*. **73**, 260-269.

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

**Weitzman, M.D., Kyöstiö, S.R.M., Kotin, R.M., and Owens, R.A.** (1994). Adeno-associated virus (AAV) rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. *Proc. Natl. Acad. Sci. USA.* **91**, 5808-5812.

**Wistuba, A., Weger, S., Kern, A., and Kleinschmidt, J.A.** (1995). Intermediates of adeno-associated virus type 2 assembly: identification of soluble complexes containing rep and cap proteins. *J. Virol*. **69**, 5311-5319.

**Wonderling, R.S., and Owens, R.A.** (1997). Binding sites for adeno-associated virus Rep proteins within the human genome. *J. Virol.* **71**, 2528-2534.

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

**Wu, J., Zhao, W., Zhong, L., Han, Z., Li, B., Ma, W., Weigel-Kelley, K.A., Warrington, K.H., and Srivastava, A**. (2007). Self-complementary recombinant adeno-associated viral vectors: packaging capacity and the role of the rep proteins in vector purity. *Hum. Gene Ther.* **18**, 171-182.

**Xiao, X., Li, J., and Samulski, R.J.** (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J Virol.* **72**, 2224–2232.

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

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

**Yamamoto, N., Suzuki, M., Kawano, M., Inoue, T., Takahashi, R., Tsukamoto, H. Enomoto, T., Yamaguchi, Y., Wada, T., and Handa, H.** (2007). Adeno-associated virus site-specific integration is regulated by TRP-185. *J. Virol*. **81**, 1990-2001.

**Yan, Z., Zak, R., Zhang, Y., and Engelhardt, J.F. (**2005). Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. *J. Virol.* **79**, 364-379.

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

**Yang, J., Zhou, W., Zhang, Y., Zidon, T., Ritchie, T., and Engelhardt, J.F**. (1999). Concatamerization of adeno-associated virus circular genomes occurs through intermolecular recombination. *J. Virol.* **73**, 9468-9477.

**Young, S.M Jr., McCarty, D.M., Degtyareva, N., and Samulski, R.J.** (2000). Roles of adeno-associated virus Rep protein and human chromosome 19 in site-specific recombination. *J. Virol.* **74**, 3953-3966.

**Zentilin, L., Marcello, A., and Giacca, M.** (2001). Involvement of cellular double-stranded DNA break binding proteins in processing of the recombinant adeno-associated virus genome. *J. Virol.* **75**, 12279-12287.

**Zhang, C., Cortez, N.G., and Berns, K.I.** (2007). Characterization of a bipartite recombinant adeno-associated viral vector for site-specific integration. *Hum. Gene Ther.* **18**, 787-797.

**Zhao, W., Wu, J., Zhong, L., and Srivastava, A.** (2006). Adeno-associated virus 2-mediated gene transfer: role of a cellular serine/threonine protein phosphatase in augmenting transduction efficiency. *Gene. Ther.* **14**, 545-550.

**Zhijian, W., Miller, E., Agbandje-McKenna, M., and Samulski, R.J.** (2006). Alpha2,3 and alpha2,6 N-linked sialic acid facilitate efficient binding and transduction by adeno-associated virus type 1 and 6. *J. Virol*. **80**, 9093-9103.

**Zhong, L., Li, B., Mah, C.S., Govindasamy, L., Agbandje-McKenna, M., Cooper, M., Herzog, R.W., Zolotukhin, I., Warrington Jr, K.H., Weigel-Van Aken, K.A., Hobbs, J.A., Zolotukhin, S., Muzyczka, N., and Srivastava, A.** (2008). Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc. Natl. Acad. Sci. USA.* **105**, 7827-7832.

**Zhong, L., Zhao, W., Wu, J., Zolotukhin, S., Govindasamy, L., Agbandje-McKenna, M., and Srivastava, A.** (2007). A dual role of EGFR protein tyrosine kinase signaling in ubiquitination of AAV 2 capsids and viral second strand DNA synthesis. *Mol. Ther.* **15**, 1323-1330.

**Zhong, L., Zhou, X., Li, Y., Qing, F., Xiao, X., Samulski, R.J., and Srivastava, A**. (2008). Single-polarity recombinant adeno-associated virus 2 vector mediated transgene expression in vitro and in vivo: mechanisms of transduction. *Mol. Ther.* **16**, 290-295.

**Zhou, X., Zeng, X., Fan, Z., Li, C., McCown, T., Samulski, R.J., and Xiao, X.** (2008). Adeno-associated virus of a single polarity DNA genome is capable of transduction in vivo. *Mol. Ther.* **16**, 494-499.

**Zincarelli, C., Soltys, S., Rengo, G., and Rabinowitz JE.** (2008). Analysis of AAV serotypes 1–9 mediated gene expression and tropism in mice after systemic injection. *Mol. Ther*. **6**, 1073-1080.

# BIOGRAPHICAL SKETCH

Shyam Daya was born in Johannesburg, South Africa at Marymount Catholic Hospital. He received early childhood training at a Montessori School in Lenasia, South Africa. He received elementary school education at Impala Crescent School in Lenasia, South Africa, at St. Francis of Assisi Convent School in Gujarat, India and at Dr. Phillips Elementary School in Orlando, Florida, USA. He received middle and high school education at Southwest Middle School and Dr. Phillips High School in Orlando, respectively.

He graduated from the Medical Academy at Dr. Phillips High School with honors and embarked on a BS in interdisciplinary studies with concentration in biochemistry and molecular biology at the University of Florida in Gainesville. After completing the BS degree in three years, he continued with the interdisciplinary Ph.D. program in medical sciences with concentration in genetics at the UF College of Medicine.

He received his first taste of research as an undergraduate in Dr. Mavis-Agbandje McKenna's laboratory, performing purification and crystallography experiments on Adeno-associated Virus Serotype 1. He continued his research on the biology of AAV under the guidance of Dr. Kenneth I. Berns, a pioneer in AAV research, who is now the Director of Genetics Institute at the University of Florida.

He wants to continue his development, and hopes to complement his research experience with clinical training, to become a physician scientist.