TRANSCRIPTION REGULATION OF ADENO-ASSOCIATED VIRUSES

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

_____________________________

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

_____________________________

by

CHAOYANG YE

Dr. David J. Pintel, Dissertation Supervisor

May 2007
The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

TRANSCRIPTION REGULATION OF ADENO-ASSOCIATED VIRUSES

presented by Chaoyang Ye,

a candidate for the degree of doctor of philosophy,

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

________________________________________
Professor David Pintel

________________________________________
Professor John Cannon

________________________________________
Professor Dongsheng Duan

________________________________________
Professor Mark Hannink

________________________________________
Professor David Setzer
ACKNOWLEDGEMENT

I would like first and foremost express my sincere thanks to my mentor, Dr. David Pintel, for his constant instruction, understanding and encouragement, which helped me not only to become a better scientist, but also to overcome the difficulties living in a foreign country. I would also like to thank Dr. Jianming Qiu, who set up a role model for the lab and provided selfless help when I started in the lab. I would also like to thank the faculty and staff in the Department of Molecular Microbiology and Immunology for their advice and support, especially the members of my committee, Dr. John Cannon, Dr. Dongsheng Duan, and in addition, Dr. David Setzer from Biological Sciences and Dr. Mark Hannink from Biochemistry. I would also like to thank Lisa Burger for her support and friendship, and my labmates Ramnath Nayak, Eun-Young Choi, David Farris, Loretta Sukhu and Richard Adeyemi for helpful discussion and friendship.

I would especially like to thank my family, my parents and my sister back in China, who not only encouraged me to become a good scientist, but more importantly taught me how to be a good person. Near or far, high or low, they are always by my side to help me move forward and pursue my dream.
TRANSCRIPTION REGULATION OF ADENO-ASSOCIATED VIRUSES

Chaoyang Ye

David J. Pintel, Dissertation Supervisor

ABSTRACT

The small genome size of adeno-associated viruses demand efficient gene regulation achieved by various strategies. This research compared these strategies from AAV2, AAV5 and bovine-AAV, and found that AAV2 capsid promoter P40 has a low basal level and is activated by its non-structural Rep in the presence of adenovirus while AAV5 capsid promoter P41 has a higher basal level and is not further activated by Rep. The difference of the activation ability between AAV2 and AAV5 Rep resides in the amino terminus of the protein. The higher basal level of P41 is at least partially contributed to the AP1 and CRE transcription factor binding sites upstream. Although BAAV, a closer relative to AAV5, has both sites, its capsid promoter activity is still low and can be further activated by both its own Rep and AAV5 Rep. This study not only enriched our understanding of the basic molecular biology of these viruses, which potentially would benefit gene therapy vector design, but also provided clue of the evolution migration between the animal and human AAVs.
# TABLE OF CONTENTS

ACKNOWLEDGEMENT.......................................................................................ii

ABSTRACT...........................................................................................................iii

LIST OF FIGURES..............................................................................................vi

CHAPTER

I. GENERAL INTRODUCTION
   A. General Biology of Paroviridae..............................................................1
   B. Genetic Organization of AAVs...............................................................4
   C. Transcription Profiles of AAVs..............................................................10

II. EFFICIENT EXPRESSION OF THE ADENO-ASSOCIATED VIRUS TYPE 5 P41 CAPSID GENE PROMOTER IN 293 CELLS DOES NOT REQUIRE REP
   A. Materials and Methods.................................................................20
   B. Results.............................................................................................24
   C. Discussion.......................................................................................43

III. UPSTREAM AP1- AND CREB-BINDING SITES CONFER HIGH BASAL ACTIVITY TO THE ADENO-ASSOCIATED VIRUS TYPE 5 CAPSID GENE PROMOTER
   A. Materials and Methods.................................................................49
   B. Results.............................................................................................54
   C. Discussion.......................................................................................73
IV. BOVINE ADENO-ASSOCIATED VIRUS EXHIBIT DISTINCT
REGULATION OF CAPSID GENE PROMOTER DESPITE
CLOSELY RELATED TO ADENO ASSOCIATED VIRUS 5

A. Materials and Methods.........................................................81
B. Results..............................................................................84
C. Discussion........................................................................103

V. ONGOING PROJECTS

V-i. Inhibition of MVM Replication by APOBEC3A

A. Material and Methods.........................................................108
B. Results..............................................................................109
C. Discussion........................................................................115

V-ii. Leader Sequence Control of Capsid Protein Expression in
Canine Parvovirus

A. Material and Methods.........................................................117
B. Results..............................................................................121
C. Discussion........................................................................125

VI. SUMMARY............................................................................126

REFERENCES............................................................................130

VITA.........................................................................................136
## LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-1</td>
<td>Members of Parvovirus family</td>
<td>3</td>
</tr>
<tr>
<td>I-2</td>
<td>Phylogenetic analysis of members of AAV</td>
<td>6</td>
</tr>
<tr>
<td>I-3</td>
<td>Genome organization of AAV2 and AAV5</td>
<td>9</td>
</tr>
<tr>
<td>I-4</td>
<td>AAV2 P5 promoter structure</td>
<td>12</td>
</tr>
<tr>
<td>I-5</td>
<td>Rep78 protein alignment between AAV2 and AAV5</td>
<td>16</td>
</tr>
<tr>
<td>II-1</td>
<td>Transcription profiles of different serotypes of AAV</td>
<td>27</td>
</tr>
<tr>
<td>II-2</td>
<td>AAV2 and AAV5 large Rep proteins have different activation properties</td>
<td>32</td>
</tr>
<tr>
<td>II-3</td>
<td>Differences in activation properties of AAV2 and AAV5 Rep reside within the amino terminus</td>
<td>36</td>
</tr>
<tr>
<td>II-4</td>
<td>(A) AAV2/AAV5 chimeras that have different transactivation properties bind the AAV2 P5 RBE similarly. (B) AAV5 Rep can competitively inhibit AAV2 Rep-dependent activation of P40</td>
<td>41</td>
</tr>
<tr>
<td>III-1</td>
<td>Core promoter comparison of AAV2 P40 and AAV5 P41</td>
<td>56</td>
</tr>
<tr>
<td>III-2</td>
<td>Upstream sequence analysis of the capsid promoters</td>
<td>59</td>
</tr>
<tr>
<td>III-3</td>
<td>Physical interaction between cellular factors and AP1 and CRE sites</td>
<td>63</td>
</tr>
<tr>
<td>III-4</td>
<td>E1A activates AAV5 P41 promoter, but the activation is not directly mediated by AP1 or CRE sites</td>
<td>67</td>
</tr>
</tbody>
</table>
III-5  AP1 and CRE enhance promoter activity in the viral RepCap background .................................................................70

III-6  Genome sequence comparison of AAV2 and AAV5-like viruses............75

IV-1  Basic transcription profile comparison between BAAV and AAV5 in response to adenovirus infection .................................................................86

IV-2  Transcription activation by AAV5 and BAAV Rep protein......................89

IV-3  Analysis of cis sequences of BAAV and AAV5 that control P41 transcription.......................................................................................93

IV-4  Upstream transcription analysis of BAAV and AAV5............................97

IV-5  Transcription analysis of linearized BAAV and AAV5 plasmids.........101

V-1  (A) Genome organization of minute virus of mouse (MVM)  
(B) Replication of MVM.....................................................................106

V-2  APOBEC3A inhibition of MVM replication is not at the ssDNA.........111

V-3  APOBEC3A does not affect MVM RNA or protein production..........114

V-4  Diagrams of the constructs used in CPV upstream sequence study......120

V-5  CPV capsid RNA and protein production regulation by different upstream sequences.................................................................................124

VI-1  Summary of capsid promoter regulation of various adeno-associated viruses.......................................................................................128
I. GENERAL INTRODUCTION

A. General Biology of Paroviridae:

The Paroviridae family includes all small, isometric, non-enveloped DNA viruses that contain linear, single-stranded genomes, usually between 4 and 6 kilobases (kb) in length. The ends of the genome, comprised of short palindromic sequences, fold back on themselves to form hairpin structures, which are essential for viral replication.

The Paroviridae family is further divided into Parovirinae and Densovirinae subfamilies. The Parovirinae subfamily is comprised of five genera: Parovirus, Dependovirus, Erythrovirus, Amdovirus and Bocavirus (the latter two being newly established), whose phylogenetic relationship is illustrated in Figure I-1. The Densovirinae contains four genera: Desovirus, Iteravirus, Brevidensovirus and Pefudensovirus. This research mainly focuses on the Dependovirus genus.

The original members of the Dependovirus genus, mainly adeno-associated viruses (AAV) are, as the name implies, dependent on helper viral functions to complete their life cycle. However, phylogenetic analysis now places autonomous goose and duck paroviruses into this genus as well due to the close similarity of their non-structural genes with AAV. Under certain conditions, however, such as the treatment of host cells with mutagens or hydroxyurea, replication of some of the dependent AAVs can be detected in the absence of helper viruses. In cultured cells normally, without helper viruses, AAV2, which is
Figure I-1. Members of Parvovirus family.

(A) Phylogenetic relationship determined by the non-structural genes of members of the Parovirinae subfamily.

(B) Phylogenetic relationship determined by the non-structural genes of members of the Densovirus and Iteravirus genera of the Densovirinae subfamily.
the prototype of AAV, can integrate site-specifically into 13.4q-qter on chromosome 19. The viral genome can be reactivated from latency upon reinfection of helper viruses. Due to the minimal host immune response to AAV infection and the diversity of new isolates, AAVs hold great potentials in human gene therapy.

Nucleotide and protein sequence alignment of the most common members of AAV (Figure 1-2) demonstrates the unique evolutionary position of AAV5. It shares only 60% similarity in nucleotide sequence with AAV2, while all the other serotypes are over 70% identical. Interestingly, AAV5 and BAAV are 77% identical in nucleotide sequence and % identical in protein sequence, indicating a much closer relationship evolutionally, which is further confirmed by gene regulation study in the third chapter of this study.

B. Genetic Organization of AAVs:

Due to the small sizes of parvoviruses, the viral genomes have to be efficiently organized to carry out all the essential functions to complete viral life cycles. Members of the parvovirus group have evolved various strategies to achieve such a goal (47, 49-52).
Figure I-2. Phylogenetic analysis of members of AAV

(A) Phylogenetic relationship between different AAVs based on nucleotide sequence alignment. Direct link between two AAVs indicates higher homology thus closer evolution relationship. AAV5 is most distantly related to other human AAVs yet closely related to bovine AAV (BAAV).

(B) Percentage of homology between different AAVs based on nucleotide sequence alignment. Same result about AAV5 can be drawn as above.
(A)

(AAV1)
AAV6
AAV2
AAV3
AAV4
AAV5
BAAV

(B)

<table>
<thead>
<tr>
<th></th>
<th>AAV1</th>
<th>AAV2</th>
<th>AAV3</th>
<th>AAV4</th>
<th>AAV5</th>
<th>AAV6</th>
<th>BAAV</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAV1</td>
<td>100</td>
<td>81</td>
<td>82</td>
<td>75</td>
<td>60</td>
<td>95</td>
<td>61</td>
</tr>
<tr>
<td>AAV2</td>
<td>100</td>
<td>82</td>
<td>74</td>
<td>60</td>
<td>83</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>AAV3</td>
<td>100</td>
<td>77</td>
<td>58</td>
<td>82</td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAV4</td>
<td>100</td>
<td>58</td>
<td>75</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAV5</td>
<td>100</td>
<td>60</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAV6</td>
<td>100</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAAV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>
As the prototype of AAV, the genetic organization of AAV2 has been extensively studied. The 4.7 kb viral genome is bracketed by two inverted terminal repeat (ITR) at the ends, which serve as the origin of viral replication. The genome is organized into three transcription units, with promoters at map units 5 (P5), 19 (P19) and 40 (P40) (Figure I-3(A)). P5 promoter generated RNA encodes for the large non-structural replication proteins (Rep78 and Rep68) while P19 transcript encodes the small non-structural Rep proteins (Rep52 and Rep40). Both transcripts undergo alternative splicing via the intron which resides in the middle of the genome and is comprised of one donor and two acceptor sites, thus generating Rep78 and Rep52 from unspliced RNA and Rep68 and Rep40 from spliced RNA. P40 promoter generates RNA encoding capsid proteins. Spliced RNA using the A1 acceptor encodes VP1 and while the RNA spliced through A2 acceptor encodes VP2 and VP3 proteins via alternative translation initiation codon.

Although genome structure-wise similar, AAV5 utilizes an interestingly different strategy to generate different transcripts (51). Similar to AAV2, it also contains three promoters, P7, P19 and P41, with P7 and P19 transcripts encoding large and small non-structural Rep proteins, respectively, and P41 generates capsid proteins (Figure I-3(B)). However, instead of alternative splicing, the P7 and P19 transcripts undergo early polyadenylation using the signal inside the intron, therefore producing only Rep78 and Rep52 P41 transcript, however, reads through the polyadenylation signal and undergoes similar aforementioned
Figure I-3. Genome organization of AAV2 and AAV5

(A) Genome structure of AAV2. AAV2 genome is shown at the bottom of the figure. Nucleotide numbers indicate the positions of ITR and three promoters, P5, P19 and P40. Positions of intron donor (D) and acceptors (A1 and A2) are also designated. The top of the figure illustrates the transcripts and protein-coding regions derived from the three promoters.

(B) Genome structure of AAV5. Similar schematic diagram of the AAV5 genome as above is shown. In addition, the proximal [(pA)p] and distal [(pA)d] polyadenylation sites are indicated. P7 and P19 transcripts terminate using the (pA)p site, while P41 mainly utilize (pA)d site.
Adapted from Qiu et al. (52)

Adapted from Qiu et al. (51)
alternative splicing as AAV2 to generate VP1 and VP2. VP3 is also made by alternative translation initiation.

Due to the high homology between AAV5 and BAAV genome sequence, similar genetic map was proposed and the usage of internal polyadenylation site has been confirmed by our recent study (49). Despite these common elements, BAAV utilizes a slightly different strategy to regulate gene expression, which is the focus of Chapter IV of this study.

C. Transcription Profiles of AAVs:

The promoter organization of the prototypic AAV2 has been comprehensively researched. In addition to the TATA box, the P5 promoter of AAV2 contains two YY1 binding sites, YY1-60 and YY1+1, the latter of which overlaps with the initiator site, and an additional major late-transcription factor (MLTF) binding site at -80 (Figure I-4). At position -20 of the promoter between TATA box and initiator resides a Rep78/68-binding element (RBE) which plays an important role in transcription control. In the absence of helper virus, P5, P19 and P40 promoters all show low levels of basal activity even in 293 cells, which partially exerts adenovirus helper function due to its constitutive expression of E1A and E1B genes. For P5, Rep binding to this RBE represses its transcription in concert with adjacent binding of cellular repressors YY1-60 and major MLTF-80 (61). With adenovirus co-infection YY1 and MLTF both activate P5
Figure I-4 AAV2 P5 promoter structure

AAV2 P5 promoter region analysis (AAV2 nucleotides 147 to 308) (J Virol. 1997 Feb;71(2):1079-88). Vertical lines and the corresponding nucleotide positions are shown on top of the sequence. Transcription factors binding sites mentioned in the text (YY1, MLTF, RBE and TATA binding protein) are indicated in boxes. +1 position indicates the site of transcription initiation.
transcription. Rep binding to RBE still represses P5 but now is able to activate downstream promoters, P19 and P40, possibly through a loop structure formed by interaction between P5-bound Rep and cellular transcription factor SP1 bound to downstream P19 and P40 promoters (7, 61).

In contrast, AAV5 P41 basal transcription is higher in 293 cells compared to AAV2 and it is not further activated upon adenovirus co-infection. Surprisingly, BAAV, the closer relative of AAV5, also exhibits activation pattern by its Rep protein in response to adenovirus when the RBE in ITR is used as the binding site (49).

The compact size of AAV demands the versatility of its gene products. The non-structural Rep protein is known to exhibit various biochemical activities and carry out vast functions during viral life cycle. For example, the bigger Rep78 and Rep68 are site-specific DNA binding proteins (12, 44, 65), as well as site-specific endonucleases (28). They also have helicase and ATPase activities (28), properties shared by Rep52 (64). A ligase activity has also been demonstrated (63). Meanwhile, many cellular proteins have been identified to bind to Rep protein due to its structural and biological flexibility. The aforementioned Rep78/68-mediated activation of P19 and P40 promoter in the presence of adenovirus involves interaction with transcription factor SP1 (30, 46). The C-terminal domains of the unspliced Rep isoform (Rep78 and Rep52) interact with cAMP-dependent protein kinase A (PKA) and its homolog PrKX, which in turn inhibit their activity. Other identified cellular Rep binding partners includes high mobility chromosomal protein HMG1 (16), TATA box binding protein TBP (24),
tumor suppressor protein p53 (3), transcriptional coactivator PC4 (71),
topoisomerase I binding RS-rich proteins Topors (70) and KCTD5, a protein
sharing homology with the cytoplasmic tetramerization domain of voltage-gated
potassium channels (69). However, the actual effect on cellular function or viral
life cycle by these interactions has not been identified.

Alignment between AAV2 and AAV5 Rep amino acid sequences shows a
high homology (Figure I-5), especially at the region where important biochemical
functions are carried out. Previous study comparing ITR binding capability of the
two proteins indicates that they are able to bind the ITRs of the other serotype.
However, the endonuclease activities are specific to their own ITR respectively.
Therefore, simple binding of the Rep protein does not ensure the proper
execution of its biochemical functions.
Figure I-5. Rep78 protein alignment between AAV2 and AAV5 using Vector NTI software (Invitrogen). Identical amino acids are shaded in dark grey and conserved amino acids are in lighter grey. Amino acids identified crucial for Rep activity are also shown.
It is interesting to investigate the molecular mechanisms behind such regulation because it provides us not only the knowledge of the diversity of viral gene expression control but also potential clues to host immigration and viral evolution. The first and second parts of this study (Chapter II and III) focuses on the differences between AAV2 and AAV5, in the aspect of viral gene activator and promoter sequence comparison, which cause the differential gene regulation. The third part (Chapter IV) compares AAV5 and BAAV, in an attempt to address their differences in capsid promoter regulation. Chapter II and III have been published in Journal of Virology (75, 76) and Chapter IV is in the process of preparation for submission.
II. EFFICIENT EXPRESSION OF THE ADENO-ASSOCIATED VIRUS TYPE 5
P41 CAPSID GENE PROMOTER IN 293 CELLS DOES NOT REQUIRE REP

Appropriate levels of capsid protein gene expression must be achieved and maintained throughout parvovirus infection. Although parvoviruses that have a single promoter (e.g. B19 and AMDV) utilize posttranscriptional mechanisms to specifically control capsid protein levels (47, 77), those parvoviruses with internal capsid gene promoters (e.g. MVM and AAV), control capsid protein accumulation by regulated transcriptional activation of these promoters (32, 52).

Transcriptional transactivation, mediated by the large non-structural proteins of these viruses (NS1 and Rep, respectively), assures the appropriate timing and expression of these promoters and hence the levels of their encoded RNA and protein products.

Transactivation of P40 by AAV2 Rep requires binding to the transcription template at either the AAV2 ITR or P5 promoter, and has been proposed to work, at least in part, by stabilizing a loop-like structure that localizes the P5 promoter, and presumably P5-associated transcription factors, to the P40 promoter (46). Efficient activation by Rep also requires co-infection by adenovirus, even in 293 cells which constitutively express the adenovirus type 5 (Ad5) E1A, a well characterized transcription activator, and E1B proteins.

AAV5, another serotype of the human AAVs, shares only 64% overall nucleotide identity with AAV2. The AAV2 and AAV5 Rep proteins share 67% amino acid identity. Both AAV2 and AAV5 Rep proteins bind the Rep-binding
element (RBE) in both AAV2 and AAV5 ITRs (10), however, the AAV2 and AAV5 Rep proteins cannot substitute for one another with respect for Rep-mediated DNA replication; because of sequence differences within their terminal resolution sites (TRS), each protein can only fully process its native origin (10).

Additionally, in contrast to AAV2, the AAV5 large Rep-gene promoter (P7) contains only a poorly-consensus RBE, suggesting that the activation of the AAV5 capsid gene promoter (P41) may be governed differently than its AAV2 counterpart.

Here we show that in contrast to other AAV serotypes and other parvoviruses with internal capsid gene promoters, the constitutive activity in 293 cells of the AAV5 P41 promoter in the ITR-deleted RepCap constructs is high, and is not significantly activated further by its large Rep protein or Ad5 infection.
MATERIALS AND METHODS

Cells and Virus:

HeLa and 293 cells were propagated as previously described (40). Transfections, using Lipofectamine and the Plus reagent (Invitrogen, Carlsbad, CA), were performed as previously described (52), and when Ad5 was co-infected, this was done 5 hours after transfection at a moi of 5.

Plasmid Constructs:

RepCap and RepStopCap plasmids: AAV2 RepCap (containing AAV2 nts 145-4492) and AAV5 RepCap (containing AAV5 nts 185-4448) have been described previously (51, 52). AAV1 RepCap (containing AAV1 nts 181-4566), AAV3 RepCap (containing AAV3 nts 181-4560), and AAV6 RepCap (containing AAV6 nts 181-4560), were constructed by inserting AAV1 and AAV3 specific sequences amplified directly from these viruses (obtained from the American Type Culture Collection), and AAV6 specific sequences amplified from the pAAV6 infectious clone (55), into the polylinker of pBluescript SK(+) (Stratagene, La Jolla, Calif.), between the EcoRI and XbaI sites. The AAV4 RepCap plasmid was a gift from J.A. Chiorini (13). The AAV2 RepStopCap and AAV5 RepStopCap constructs have premature termination codons introduced at nts 489 and 480, respectively, in the amino terminus of their respective Rep proteins. To create the AAV1, AAV3, AAV4, and AAV6 RepStopCap plasmids, frame shift mutations were introduced in the parent RepCap plasmids within the large Rep coding region at nts 743, 523, 781, and 729, respectively, leading to termination of these proteins shortly downstream.
P40 and P41 minimal transcription unit test plasmids: P40VP contains AAV2 nts 1700-4492. P41Cap contains AAV5 nts 1637-4448. Sequences from AAV2 P5 (nt 146 to 320) (52), AAV5 P7 (nt 168-358) (51), the AAV2 ITR (nt 1-145) (52), or the AAV5 ITR (nt 1-183) (51), were inserted into P40VP at nt 1700 (123 nts upstream of the P40 TATA box), or into P41VP at nt 1637 (243 nts upstream of the P41 TATA box), to create P5P40VP, P7P40VP, V2ITRP40VP, V5ITRP40VP, or P5P41VP, P7P41VP, V2ITRP41VP, and V5ITRP41VP, respectively, as shown in the diagram in Figure II-2.

pHelper and Rep expressing plasmids: pHelper plasmid is from Stratagene (TX). Rep expression constructs were driven by the HIV promoter within the HIV long terminal repeat promoter and have been described previously (51, 52).

AAV2/AAV5 chimeric Rep plasmids and the P5P40 luciferase reporter plasmid: Chimeric Rep constructs were based on the HIV-driven, AAV2 and AAV5 Rep-expressing parental constructs previously described (51, 52). The numbers in the name of each construct designate exactly where the chimeric borders were made. All fusions were made in frame and sequenced to ensure they were as predicted. The V2 211(V5) series were based on the V2 211 parental construct. To generate the P5P40 luciferase construct, the AAV2 P5 promoter (nt 146 to 320 from AAV2) was cloned upstream of the luciferase gene driven by AAV2 P40 (nts 1700-1883), to provide an RBE, but was cloned in an inverted orientation relative to the direction of P40 transcription to exclude potential P5-generated luciferase activity.
*Chimeric Rep binding domain plasmids and the P5P40 minimal transcription unit test plasmid:* The parental RepBD.TZ.AD and RepBD.TZ plasmids were gifts from Matt Weitzman (Salk Institute, San Diego, Calif.), and contain a modified leucine zipper dimerization domain from GCN4 (TZ), and in the case of RepBD.TZ.AD, the activation domain from herpes virus VP16 (6). RepBD.TZ.AD- and RepBD.TZ-derived constructs containing either AAV5 or chimeric Rep sequences (V5 101-121, V5 121-161 and V5 161-181) were constructed by replacing the amino terminal 244 aa Rep-binding domain from the original AAV2 binding domain construct with either AAV5 or chimeric Rep sequences. The P5P40 minimal transcription unit test construct is the same as described above.

**RNase Protection Assays:**

Total RNA was isolated 36-41 hours post-transfection as previously described (42, 57). RNase protections were performed as previously described (42, 57), using homologous anti-sense probes, termed “RP”, which spanned nts 1781-1923 (for AAV1), 1767-1906 (for AAV2), 1764-1903 (for AAV3), 1824-1957 (for AAV4), and 1766-1908 (for AAV6), of the individual P40 promoters; and nts 1846-1985 of the P41 promoter of AAV5, respectively.

**Luciferase Assays:**

Luciferase assays were performed according to (32). Briefly, 293 cells grown in 12 well plates were transfected with the P5P40 luciferase reporter construct together with pHHelper and the Rep expression constructs described in the text at a 1:4:0.25 ratio. In addition, 0.05 ug per well of CMV driven β-
galactosidase reporter was co-transfected as internal control. The total amount of DNA transfected into each well was kept at 0.6 ug. 36 hours after transfection, cells were lysed with lysis buffer containing 1\% Triton X-100. Equal amounts of the sample were used to test luciferase activity and \( \beta \)-galactosidase activity according to the manufacturer's instructions (Applied Biosystems (Tropix), Bedford, MA). Each experiment represents the average of duplicates from three individual experiments (error bars shown), each normalized to \( \beta \)-galactosidase activity.

**Rep inhibition assays:**

Activity of the P5P40 luciferase construct in response to an HIV-driven AAV2 Rep plasmid was first assayed (data not shown), and the lowest amount of Rep-expressing plasmid giving the highest level of activity was determined. Luciferase activity was then monitored in assays in which the total DNA concentration was kept constant, following addition of decreasing amounts of the HIV-driven AAV2 Rep expressing plasmid, using either empty vector, or the HIV-driven AAV5 Rep-expressing plasmid to make up the balance. Each experiment represents the average of duplicates from three individual experiments (error bars shown), each normalized \( \beta \)-galactosidase activity.
RESULTS

*The constitutive activity in 293 cells of AAV5 P41 in RepCap constructs is high, and is not significantly activated further by its large Rep protein or adenovirus infection.*

Maximal levels of activation of AAV2 P40, in the presence of Ad, occur from full-length, ITR-containing clones. While the ITR can supply the RBE required for Rep activation of AAV2 P40, it is likely that these maximal levels of activation are partially attributable to amplification of templates due to replication, because, as mentioned above, in the presence of adenovirus, the AAV2 P5 RBE can fully support P40 activation in non-replicating AAV2 RepCap constructs (52). Thus, for our analyses, we chose to utilize ITR-deleted Rep-Cap constructs to examine the effect of Ad and Rep on the activation of the capsid gene promoters of various AAV serotypes, thereby eliminating expression differences due to variation in viral replication under these different conditions. Although the AAV5 P7 promoter does not contain a consensus RBE, the AAV5 RepCap construct could be included in this analysis because the level of RNA generated from the AAV5 capsid gene promoter (P41), in an ITR-lacking AAV5 RepCap clone, was surprisingly high (51).

As shown in Figure II-1A, expression, in 293 cells, of both the upstream P5 and P19 promoters and the P40 promoters within Rep-expressing RepCap constructs of various AAV serotypes, was high, both in the absence and presence of Ad5. However, the P40 promoters of AAV1, 2, 3, 4 and 6 were all activated by Ad5 significantly more than was the P41 promoter of AAV5 (Figure
Figure II-1. Transcription profiles of different serotypes of AAV. (A). RNase protection assay, performed as previously described (42, 57), using homologous RP probes (spanning from approximately 86 nts upstream to 54 nts downstream of each P40, or 66 nts upstream of and 73 nts downstream of the P41, initiation site) to protect RNA generated in 293 cells following transfection of RepCap constructs of different serotypes of AAV in the absence (-) or presence (+) of adenovirus infection. Bands protected by transcripts from upstream P5 (P7 for AAV5), P19 promoters, or P40 (P41 for AAV5) promoters are indicated. Multiple bands for P40 transcripts represent alternative initiation events and some probe chatter; all bands were included in quantifications. (B). This gel shows the same assay performed on RNA generated by transfection of RepStopCap constructs. The AAV2 RepStopCap and AAV5 RepStopCap constructs have premature termination codons introduced at nts 489 and 480, respectively, in the amino terminus of their respective Rep proteins. The AAV1, AAV3, AAV4, and AAV6 RepStopCap constructs contain frame-shift mutations in the amino terminus of their respective Rep proteins at nt 743, 523, 781, and 729, respectively, which lead to termination of these proteins shortly downstream. (C). Quantification, using a Molecular Imager FX and the Quantity One version 4.2.2 image software (Bio-Rad, Hercules, CA), of RNase protections, a representative of which is shown in panel A. Data from at least three experiments, with standard error bars, are presented as the activation fold of P40(P41) promoters comparing transcription levels in the presence or absence of adenovirus. (D). Comprehensive quantification of RNase protections of the type shown in panel A
and B. P40 (P41) transcription levels are calculated as the percentage of total transcripts (P5+P19+P40). Quantifications of at least three individual experiments, including standard error bars, are shown. (E). RNase protection assay with the same probes described in part A. of RNA generated in HeLa cells following transfection of RepCap and RepStopCap constructs of AAV2 and AAV5. Bands protected by transcripts from upstream P5 (P7 for AAV5) and P19 promoters, or P40 (P41 for AAV5) are indicated. (F). Sequence alignment of AAV2 P40 (nt 1690-1908) and AAV5 P41 (nt 1726-1991) promoters. Consensus nucleotides are shaded. TATA boxes are indicated. SP1 and GGT elements previously reported to mediate AAV2 P40 activation (46) are also indicated.
II-1A, compare lanes 1-8 and 11-12 to lanes 9-10), both as measured directly (Figure II-1C), and as a percentage of total RNA [Figure II-1D, compare open bars (RepCap (-) Ad5) to black bars (RepCap (+) Ad5). Identical amounts of RNA were used for each protection. Because no suitable internal transfection control, whose expression was independent of Ad5, was identified, direct activation levels (Figure II-1C) were calculated as the average of at least three separate complete experiments that included all test plasmids; and various plasmid preparations were tested. Southern analysis was initially used to detect and standardize for cell-associated plasmid DNA after transfection as previously reported (77). However, for our experiments, this was not proportional to DNA that had entered the nucleus (data not shown), and so was also deemed an unsuitable control for these purposes. As previously reported (52), in 293 cells the activity of the P40 promoters in RepCap constructs remained relatively low compared to the activity of the P5+P19 promoters, even in the presence of Ad5, and was significantly lower than seen during AAV2 infection.

That expression of the AAV5 P41 promoter was high in these assays was surprising because the AAV5 RepCap plasmid is not predicted to contain a Rep binding site. This suggested that either AAV5 Rep activation of P41 did not require binding to the transcription template, or that P41 activity was independent of Rep under these conditions. The latter seemed to have been the case, because in contrast to the other AAV serotypes, the activity of the P41 promoter also remained high in similar constructs in which expression of the Rep protein was abolished by the introduction of a translation termination signal at nt 480 in
the amino terminus of the gene (RepStopCap plasmids), both in the presence or absence of Ad5 (Figure II-1B, compare lanes 13-20 and 23-24 to lanes 21-22; Figure II-1D). The addition of homologous Rep proteins in trans to AAV1, 2, 3, 4, and 6 RepStopCap constructs enhanced their P40 promoters to various extents, however, AAV5 P41 was not further activated (data not shown). These results suggested that basal activity of P41 in 293 cells was constitutively high and not further enhanced by either Ad5 infection or Rep.

Expression of AAV5 P41 was also independent of the AAV5 Rep protein in HeLa cells, and in these cells, the difference between AAV5 P41 and AAV2 P40 was even more marked. In HeLa cells the basal levels of both the upstream promoters and the capsid–gene promoters of both AAV2 and AAV5 RepCap constructs were considerably less than seen in 293 cells (Figure II-1E, lanes 1 and 5, respectively), and the levels of both of these promoters were thus dramatically enhanced by adenovirus (Figure II-1E, lanes 2 and 6, respectively). Full levels of activation for the AAV2 promoters in HeLa cells were dependent upon the Rep protein (Figure II-1E, compare AAV2 RepCap, lanes 1 and 2, to AAV2 RepStopCap, lanes 3 and 4), however, as in 293 cells, activity of the AAV5 constructs was independent of Rep expression (Figure II-1E, compare AAV5 RepCap, lanes 5 and 6, to AAV5 RepStopCap, lanes 7 and 8). Furthermore, these results suggested that the higher activity of AAV5 P41 in 293 cells shown in Figure II-1B was likely due to the resident Ad E1A and or E1B gene products. Thus, while expression of the AAV2 P40 promoter requires Ad5 infection plus
Rep in 293 cells, expression of AAV5 P41 can be sustained by E1A and E1B alone.

When compared at optimal alignment there are significant differences between the two promoters that may account for the higher basal activity of P41 (Figure II-1F). Interestingly, the Sp1-50 binding site, which is important for AAV2 Rep activation of P40 (46), is significantly altered in P41. A comprehensive comparison of the two promoters is in progress.

Although the P41 promoter did not require Rep for activation, because the AAV5 RepCap constructs are predicted to lack an AAV5 Rep binding site, the intrinsic transcription-transactivating potential of the AAV5 Rep protein needed to be tested in another way.

**In contrast to the AAV2 Rep protein, the AAV5 Rep protein is a poor transactivator of the inducible AAV2 P40 promoter.**

To examine the activation potential of AAV5 Rep further, we chose to examine transactivation of the AAV2 P40 promoter resident within a series of minimal constructs to which various naturally occurring potential Rep binding sites were attached. As expected, the AAV5 P7 promoter region, which lacks a consensus binding site for either the AAV2 or AAV5 Rep proteins, conferred to P40 only a modest amount of activation by either the AAV2 or AAV5 Rep proteins (Figure II- 2A, P7P40VP, lanes 7-9), compared to the background parent P40VP (Figure II-2A, lane 1-3). When a Rep-binding element was supplied by the AAV2 P5 promoter (P5P40VP), activation by AAV2 Rep increased to
Figure II-2. AAV2 and AAV5 large Rep proteins have different activation properties. RNase protection assay, performed as previously described (42, 57), using a homologous RNase protection probe spanning either the P40 (nt 1767-1906) or P41 (nt 1846-1986) initiation site (designated RP in the diagrams at the top of each panel), of RNA generated in 293 cells following transfection of P40VP (nt 1700 to 4492) (A), or P41Cap (nt 1637 to 4448), (B), linked with different upstream Rep-binding sequences (diagramed on the top of each gel).

pBluescript SK+ was used as empty vector control. AAV2 and AAV5 Rep were provided in trans and driven by HIV promoter as previously described (51, 52). All transfections were followed by adenovirus infection to allow activation.

Sequences from AAV2 P5 (nt 146 to 320) (52), AAV5 P7 (nt 168-358) (51), the AAV2 ITR (nt 1-145) (52), or the AAV5 ITR (nt 1-183) (51), were inserted in front of P40VP or P41VP (as shown in the diagram), to create P5P40VP, P7P40VP, V2ITRP40VP, V5ITRP40VP, or P5P41VP, P7P41VP, V2ITRP41VP, and V5ITRP41VP, respectively. The human β-actin probe was used as a loading control. The arrowhead indicates transcripts generated from upstream promoter activities. The asterisk indicates non-specific band associated with β-actin protection. The activation fold generated by AAV2 or AAV5 Rep was calculated by setting empty vector as 1.
approximately 7-fold over background levels (Figure II-2A, compare lanes 4 and 5), however, activation by AAV5 Rep remained at approximately 2-fold (Figure II-2A, compare lanes 4 and 6, respectively). When the AAV2 ITR was used to supply the RBE (V2ITRP40VP), transactivation by AAV2 Rep, but not by AAV5 Rep, was enhanced further yet (Figure II-2A, compare lane 10 to 11 and 12, respectively). The AAV5 Rep protein exhibited somewhat higher activation of AAV2 P40 when the AAV5 ITR was linked in cis (V5ITRP40VP), approximately 3-4 fold over background, similar to levels achieved by AAV2 Rep with this construct (Figure II-2A, compare lane 13 to lanes 14 and 15). The AAV5 and AAV2 Rep proteins accumulated to similarly high amounts following transfection of 293 cells as detected by western-blots analysis (data not shown), and the AAV5 Rep protein was also shown to bind these elements as well as did the AAV2 Rep, as measured by the ability to target the VP16 activator to a test template (see below).

These results suggested that the AAV5 Rep protein was a significantly less potent activator of the inducible P40 promoter than was AAV2, even when targeted to the transcription template. As expected, the AAV5 P41 promoter was only poorly inducible by either AAV2 or AAV5 Rep, regardless of the nature of the RBE linked in cis, primarily because of its higher basal activity (Fig 2B).

* A difference in the activation potential between AAV2 and AAV5 Rep can be mapped to the N-terminal domains of the molecules. *
Although AAV2 and AAV5 Rep proteins share significant sequence homology, and can both bind to an (GAGY)₃ consensus binding site with similar affinities in vitro (10), they activated the inducible AAV2 P40 promoter to significantly different levels when targeted to the transcription template by either the AAV2 P5 or ITR RBE. To determine regions of the two Rep proteins that might distinguish their transactivation properties, we made a series of chimeras between the two (diagrammed in Figure II-3A), and tested them for their ability to activate the inducible AAV2 P40 promoter. Constructs expressing luciferase from a P40 promoter with an RBE-containing AAV2 P5 element linked upstream in an orientation inverted relative to the direction of transcription of P40 (Figure II-3B, top), were co-transfected in 293 cells with the pHHelper plasmid (Invitrogen), which supplies Ad5 E2A, E4orf6 and VA RNA. All of the wild-type and chimeric Rep proteins assayed in the following experiments accumulated to similar levels as assayed by Western blot (data not shown).

Consistent with results presented in the previous sections, the full-length AAV2 Rep protein was a strong activator in this system (V2, Figure II-3B). The full-length AAV5 Rep protein, however, which bound the AAV2 P5 RBE to an extent similar to that of AAV2 Rep in a heterologous binding assay described below, showed dramatically reduced activation potential (V5, Figure II-3B). Analysis of chimeras in which carboxyl-terminal residues of two proteins were exchanged demonstrated that the amino terminal region of these Rep proteins governed their transactivation properties in these assays. Both V2 362 and V2 211, which contained the AAV2 Rep amino terminal domain retained
Figure II-3. Differences in activation properties of AAV2 and AAV5 Rep reside within the amino terminus. (A). Illustration of the chimeric Rep constructs described in the text. Filled bars indicate sequences derived from AAV2 Rep. Open bars indicate sequences derived from AAV5 Rep. The numbers in the name of each construct designate exactly where the chimeric borders were made. All fusions were made in frame and sequenced to ensure they were as predicted. All Rep expression was driven by the HIV promoter, in constructs similar to those previously described (51, 52). The V2 211(V5) series were based on the V2 211 parent construct. (B). Top, diagram of the P5P40 luciferase reporter plasmid used for these assays as described in the text. The AAV2 P5 promoter (nt 146 to 320 from AAV2) was cloned upstream of the P40 (nt 1700-1883) luciferase gene, to provide an RBE, but was cloned in an inverted orientation relative to the direction of P40 transcription to prevent P5-generated luciferase activity. pHelper plasmid was co-transfected to provide helper adenovirus gene products. Bottom: Quantification of wild-type and chimeric Rep transactivation. Activity of each protein is shown with standard error bars, and is the average of duplicate samples from at least three experiments. All values are standardized to β-galactosidase activity generated by a CMV driven β-galactosidase expression vector used as internal control.
transactivation activity similar to AAV2 Rep, while V5 362 and V5 211, which contained the amino terminal region of AAV5 Rep, were as inactive in this assay as the wild-type AAV5 Rep (Figure II-3B). Based on these results, chimeras within the first 211 amino acids were generated within the AAV5 Rep background (within the V2 211 parent) to attempt to identify regions which were responsible for differences between the activation properties of these proteins. These chimeras exhibited variability in their ability to transactivate the test template. V2 211-generated Rep proteins, in which either aa 181-211 or aa 161-181 was replaced with residues of AAV5 Rep, retained approximately half of the transactivation activity of the parent protein [V2 211(V5 181-211) and V2 211(V5 161-181), respectively, Figure II-3B]. Replacement of V2 211 amino acids 101-121 with those of AAV5 reduced the activation activity of the molecule further [V2 211(V5 101-121), Figure II-3B], while replacement of amino acids 1-57, 57-101 or 121-161, reduced activity to the very low levels seen for the AAV5 wild-type Rep molecule [V2 211(V5 1-57), V2 211(V5 57-101), and V2 211(V5 121-161), respectively, Figure II-3B]. These results suggested that a region that lies within the previously identified DNA binding region of AAV2 Rep governs its activation properties. While the full 211 amino termini of these Rep proteins are important, these results generally suggest that the amino-terminal 100 amino acids are most critical for activity.

Although the chimeras were made from two parents each competent for binding to an RBE, three amino-terminal chimeras were further analyzed to reveal any potential differences in their ability to bind to the P5 RBE which might
account for differences in their activation activities (Figure II-4). This was done using a heterologous reporter assay, in which DNA binding was measured indirectly via targeting of the VP16 activation domain. The amino terminal 244 aa Rep binding domains from the wild-type or chimeric Rep proteins assayed above, fused with a modified leucine zipper dimerization domain from GCN4 [T.Z., (6)] and the VP16 activation domain, were assayed for levels of VP16-dependent activation of a P5P40 test plasmid (diagrammed in Figure II-4A, top). As expected from previously published reports (10), the AAV2 and AAV5 amino-terminal binding domains (V2 BD.TZ.AD and V5 BD.TZ.AD, respectively), exhibited similar levels of binding to the P5 RBE (as measured by VP16 activation of P40, Figure II-4A, compare lanes 3-4 to 5-6). Three of the chimeras tested for activation above, V5 161-181 BD.TZ.AD, which exhibited the highest level of activation, V5 101-121 BD.TZ.AD, which exhibited an intermediate level of activation, and V5 121-161 BD.TZ.AD., which exhibited little if any activation, all exhibited similar levels of binding to the P5 RBE as measured in this assay (Figure II-4A, compare lanes 7-8, 9-10 and 11-12, respectively). Thus, the chimeras bound to the AAV2 P5 RBE as well as the two parents and gross differences in binding could not explain differences in their activation properties. Activity over vehicle alone transfection (SK), was not detected following transfection of a Gal4VP16 chimera (Gal-VP) (Figure II-4A, compare lanes 1 and 2), or when the VP16 activation domain was not present on the Rep binding domain/GCN4 leucine zipper domain fusions (Figure II-4A, compare lane 2 to lanes 3, 5, 7, 9, and 11). Activation of the P5 promoter by VP16 showed a
Figure II-4. (A). AAV2/AAV5 chimeras that have different transactivation properties bind the AAV2 P5 RBE similarly. The P5P40VP reporter plasmid and the RNase protection probe used in this experiment are diagramed. RNase protection assays were performed as previously described (42, 57), using the RP probe described above (as shown in the diagram) on RNA generated in 293 cells following co-transfection of the P5P40VP construct with different Rep chimeras as described in the text, and as indicated above each lane of the gel. Protection by the human β-actin probe was used as a control. The arrowhead indicates transcripts generated from upstream promoter activity. The asterisk indicates non-specific band associated with β-actin protection. For quantification, all P40 transcription levels were normalized to β-actin levels, and the activity of each fusion protein was reported relative to P40 activity in response to the pSK+ vehicle (lane 1) which was set to 1. As described in the text, wild-type and chimeric amino terminal fusions were assayed with and without the VP16 activation domain. (B). AAV5 Rep can competitively inhibit AAV2 Rep-dependent activation of P40. The same luciferase reporter assay as described in Figure II-3 was used in this experiment. Activity in response to an HIV driven AAV2 Rep plasmid was first assayed. The lowest amount of plasmid giving the highest level of activity was determined. These activities (filled bar), as well as activity in response to co-transfection of the parent pSK+ (open bar), are shown on the right. In the samples on the left, decreasing amounts of AAV2 expressing plasmid (the fraction of which is indicated on the bottom of each pair) was mixed, keeping the final DNA concentration the same, with either empty vector (open
bars), or a plasmid expressing AAV5 Rep from an HIV promoter (filled bars). At each amount of AAV2 Rep-expressing plasmid used, the co-transfected excess AAV5 Rep-expressing plasmid reduced AAV2 Rep-dependent activation greater than did equal amounts of co-transfected vehicle. Each experiment represents the average of duplicates from three individual experiments (error bars shown), each normalized to β-galactosidase activity expressed from a co-transfected expression vector in which β-galactosidase was driven by a CMV promoter.
similar profile (Figure II-4A, arrowhead). All Rep fusion proteins were expressed at similar levels in these assays (data not shown). These results suggested that the transactivation potential of the two Rep proteins could be separated from their binding activity, at least in these simple binding assays.

Since both the AAV2 and AAV5 Rep proteins bound to the AAV2 P5 RBE, yet show significant differences in activation activity, we tested whether the AAV5 Rep protein could compete for, and thus inhibit AAV2 Rep activation of a P5 RBE-linked reporter template. As seen in Figure II-4B, the addition of excess AAV5 Rep-expressing plasmid, but not the plasmid vehicle, inhibited activation of these reporter templates in response to a broad range of AAV2 Rep-expressing plasmids.
DISCUSSION

Parvoviruses generate capsid-encoding mRNAs either from an internal promoter (32, 52), or by RNA processing of a large pre-mRNA generated from a single promoter (47, 77). In previously characterized examples, expression from internal capsid gene promoters is low in all cell types examined prior to activation either by a viral protein alone (the virus large non-structural protein for the Parvovirus genera and GPV), or by the viral Rep protein plus helper virus (for the AAV group of viruses). In this manuscript, we show that within RepCap constructs lacking ITR sequences, expression in 293 cells of the AAV5 P41 capsid gene promoter is high, and is not further activated by either its large Rep protein or Ad5 infection. In addition, the AAV5 Rep protein is itself a poor activator of the inducible AAV2 P40 promoter. The sequences immediately upstream of the P41 and P40 promoters are quite different, and this may play a role in the different activity of these promoters in minimal constructs.

Does this imply that, unlike other paroviruses, expression of the AAV5 capsid proteins is unregulated during infection? It may be that even though P41 basal activity is high within minimal expression constructs in 293 cells, capsid expression from the viral genome may be regulated during infection in response to Rep and the presence of helper-virus. This is currently being examined. High basal activity of P41 requires Ad E1A and/or E1B suggesting that helper virus gene products, and perhaps Rep, play a role in P41 expression during infection.
Our results also suggest that efficient transcriptional activation by AAV2 Rep is likely dependent on the nature of its interaction with the RBE binding site in vivo. Surprisingly, AAV2 Rep-activation of the V5ITRP40VP construct was less than for those containing an authentic AAV2 RBE. Although AAV2 Rep binds both of these sites with similar affinities in in vitro gel shift assays (10), AAV2 Rep binding to the AAV5 ITR is not sufficient to fully activate the AAV2 P40 promoter. Similarly, a synthetic minimal RBE site alone is not sufficient to allow Rep-mediated transactivation of the AAV2 P40 promoter; flanking sequences including the TRS are required for optimal transactivation (data not shown and (52)). In addition, AAV2 and AAV5 Rep proteins cannot substitute for one another during DNA replication, because the TRS sequences of the AAV2 and AAV5 ITRs are functionally different (10). Thus, the nature of AAV2 Rep or AAV5 Rep binding to their respective ITRs in vivo is different, and this may account for their differences in activation potential. Alternatively, these binding sites may themselves differ in their ability to bind additional cellular factors required for the activation process.

Our examination of chimeras between the AAV2 and AAV5 Rep proteins has revealed a separation between the RBE binding and transactivation capabilities of Rep. Amino-terminal chimeras which transactivated strongly, intermediately, or weakly, showed indistinguishable binding to the AAV2 P5 RBE in heterologous binding assays. Thus, while the amino terminus of both AAV2 and AAV5 Rep are interchangeable in regards to binding of the RBE in our
assay, the amino terminus of AAV2 Rep contains regions required for transactivation that are lacking in AAV5 Rep.

An independent activation domain of Rep has not yet been identified, however, the DNA binding domain of Rep has been localized to the amino terminal region (36, 45, 72, 74), and the structure of the AAV5 Rep DNA binding domain (aa 1-197) has been determined (26). This region of AAV5 Rep protein displays a ferredoxin-like βαββαβ fold, which is topologically homologous to numerous other DNA binding proteins, most notably simian virus 40 (SV40) T antigen and the DNA binding domain of the replication initiation protein E1 of bovine papillomavirus (BPV). Interestingly, the core elements responsible for this high homology (β4, β1, β3, β2, αB and αC) (26), reside in the region of AAV5 Rep that, when transferred to the analogous region of AAV2 Rep, most significantly impaired its ability to activate in our assays (1-57: β1; 57-101: αB, β2 and β3; 101-121: αC; 121-161: β4).

Furthermore, the structure of AAV5 Rep aa 1-197 complexed with the AAV5 ITR RBE has also been determined (27). Seven amino acids that make contact with the RBE are conserved between the two proteins (27), yet they fall into the same region (aa 121-161) which our assays identify as functionally different between the two with regard to transactivation. However, differences between the two proteins in this region - specifically that AAV2 Rep has a single Arg at residue 138 rather than the Lys137 and Lys138 residues of AAV5 Rep, and that for AAV5 the downstream Gly139 and Gly140 are followed by Ala141, rather than the three sequential glycines found for AAV2 Rep - have led to the
prediction that the interaction of AAV5 with its RBE may be qualitatively different than the interaction of the other AAV Rep proteins with their RBE (26). Thus, while the conserved similarities between the AAV2 and AAV5 Rep binding domains may help explain why the two proteins behave comparably in our simple binding assays, it will be important to determine whether differences within the region required for transactivation which have been predicted to result in subtly different DNA binding properties accounts for differences in the transactivating activities of the two Rep proteins.
III. UPSTREAM AP1- AND CREB-BINDING SITES CONFER HIGH BASAL ACTIVITY TO THE ADENO-ASSOCIATED VIRUS TYPE 5 CAPSID GENE PROMOTER

Members of the Parvovirus and Dependovirus genera use internal promoters to regulate capsid gene expression (17, 53). Activity of these promoters is typically low in all cell types examined prior to activation either by a viral protein alone (e.g., the large nonstructural viral protein NS1 for the Parvovirus genera, or, in the case of the adeno-associated virus type 2 (AAV2) -like viruses, by the viral Rep protein plus helper virus (5, 17, 53). For the prototype AAV2, transactivation of P40 capsid protein requires binding of the large Rep proteins to the transcription template at either the AAV2 inverted terminal repeat (ITR) or the P5 promoter (35) and has been proposed to work, at least in part, by stabilizing a loop-like structure that localizes the P5 promoter, and presumably P5-associated transcription factors, to the P40 promoter (30, 46). Efficient activation of AAV2 P40 by AAV2 Rep requires co-infection by adenovirus (Ad), even in 293 cells which constitutively express both the well-characterized transcription activator E1A, and E1B proteins (5, 52, 67).

AAV5, the most divergent of the AAV serotypes, shares only 64% overall nucleotide identity with AAV2 (2, 10, 76). The AAV2 and AAV5 Rep proteins share 67% identity (11, 76). Both AAV2 and AAV5 Rep proteins bind the Rep-binding element (RBE) that is present in both AAV2 and AAV5 ITRs (10) and the AAV2 P5 promoter (76). In contrast to AAV2, however, the AAV5 large Rep
gene promoter (P7) contains only a poorly-consensus RBE (76), which suggested that activation of the AAV5 capsid gene promoter P41 may be governed differently from its AAV2 counterpart (76).

We have recently shown that within RepCap constructs lacking ITR sequences, the expression level of the AAV5 P41 capsid gene promoter in 293 cells was surprisingly high, and this higher level of activity required adenovirus E1A and/or E1B (51, 76). However, in contrast to AAV2 P40, neither adenovirus infection nor the large Rep protein was required to achieve these higher levels of expression in these cells (51, 76). In addition, the AAV5 Rep protein was itself a poor activator of the inducible AAV2 P40 promoter (76).

In this manuscript we show that the higher basal activity of AAV5 P41 relative to AAV2 P40 is due to differences in the relative strengths of the core promoter sequences, but more importantly, to the presence of active binding sites for the transcription factors CREB and AP1 within the upstream region of AAV5 P41 that are absent from AAV2 P40. These differences also govern the relative basal activity of the AAV5 capsid gene promoters within near-full length viral genomes.
MATERIALS AND METHODS

Cells and Virus:

HeLa and 293 cells were propagated as previously described (40). Transfections, using Lipofectamine and the Plus reagent (Invitrogen, Carlsbad, CA), were performed as previously described (52), and when Ad5 was co-infected, this was done 5 hours after transfection at a moi of 5.

Plasmid Constructs:

Luciferase reporter constructs: All luciferase reporter constructs are based on pGL3-Basic (Promega, WI). The core promoters (CP40, CP41) and the mutant sequences shown Figure III-1A were inserted between Xho I and Hind III sites in the polylinker region. Constructs shown in Figure III-2A, designed to analyze upstream sequences, were based on CP40 and CP41 constructs; upstream sequences were inserted between the Nhe I and Xho I sites. 2L, 2S and V2min span from nt 1655, 1708, and 1744, to nt 1822, respectively. 5L, 5S and V5min span from nt 1681, 1733, and 1770, to nt 1879, respectively. V2min AP1 is based on V2min with nt 1761-1766 GAGCAT changed to ACTCAC. V2min CRE is based on V2min with extra CRE-binding sequence from AV5 CTGGGTGACGTCACCAATACTAGC added at the 3’. V2min AP1 CRE contains both changes mentioned above. V5min AP1m is based on V5min with the sequence ACTCAC at nt 1787-1792 changed to GAGCAT. V5min CREm is based on V5min with the GT at nt 1865-1866 changed to TC. V5min AP1m CREm contains both changes.
**E1A, E1B expression constructs:** E1A, E1B, 13SE1A, 12SE1A expression constructs were gifts from Greg Tullis (Boston University). Based on the 13SE1A plasmid, dlN has a deletion of the N-terminal 40 amino acids of E1A, and dlCR1 has deletion of E1A aa 41-80 generated similarly to mutants characterized in (73)). All the wild type and mutant E1A expression constructs routinely generated similar levels of protein following transfection, as determined by western blot analysis.

**RepStopCapTR and RepStopCapDMTR plasmids:** The AAV2 RepStopCapTR and AAV5 RepStopCapTR constructs are based on AAV2 RepCap (containing AAV2 nts 145-4492) (52) and AAV5 RepCap (containing AAV5 nts 185-4448) (51) constructs, and contain premature termination codons introduced at nts 489 and 480, respectively, in the amino terminus of their respective Rep proteins, as previously described (76). In addition, half ITR sequence from AAV2 (nt 4489-4626) or AAV5 (nt 4476-4587), respectively, is added at the 3’ of the genome to allow Rep binding and activation. The AAV2RepStopCapDMTR and AAV5RepStopCapDMTR plasmids contain the same promoter upstream region as the reporter constructs V2min AP1 CRE and V5min AP1mCREm, respectively, as described above.

**Luciferase Assays:**

Luciferase assays were performed according to the manufacturer’s suggested protocols (Promega, WI). Briefly, 293 cells grown in 12 well plates were transfected with 0.1 ug per well of the luciferase reporter constructs along with 0.35 ug per well of empty pBluescript SK(+) (Stratagene, La Jolla, Calif.)
using the lipofectamine reagent (Invitrogen). In addition, 0.05 ug per well of TK driven renilla luciferase gene reporter was co-transfected as internal control. The total amount of DNA transfected into each well was kept at 0.5 ug. 36 hours after transfection, cells were lysed and the luciferase activity was tested using the Dual-Luciferase Reporter Assay system (Promega, WI). Each experiment represents the average of duplicates from three individual experiments (error bars shown), each normalized to renilla luciferase activity.

Expression of the TK driven renilla luciferase gene varied in response to E1A. Therefore, renilla luciferase gene activity was not a suitable as internal control when testing E1A function in HeLa cells. Thus, these experiments were carried without internal control but repeated multiple times. Using the lipofectamine reagent (Invitrogen, Carlsbad, CA), HeLa cells grown in 6 well plates were transfected with 0.3 ug per well luciferase reporter constructs together with 0.6 ug per well of empty pBluescript SK(+) DNA or different E1A constructs. In addition, 0.6 ug E1B was added in the E1A E1B cotransfection group. Otherwise, 0.6 ug pBluescript SK(+) DNA was added to make up the amount of total DNA. 36 hours after transfection, 1/5 of the cells were lysed with buffer containing 1% Triton X-100 and subsequently tested for luciferase activity as previously described (32). 4/5 of the cells were used to assay E1A expression levels by western blot analysis.

**Nuclear Extract Preparation:**

Nuclear extract preparation was adapted from Schreiber et al (58). Briefly, 293 cells were collected and washed with PBS. Cells were then resuspended in
2 volumes of Buffer A (10 mM Hepes, 10 mM KCl, 1.5 mM MgCl$_2$ and 1.0 mM DTT, pH7.9), lysed with a tight Dounce homogenizer and centrifuged at low speed to collect nuclei. Pellets were resuspended in 2 volumes of Buffer C (20 mM Hepes, 1.5 mM MgCl$_2$, 1.0 mM DTT, 0.2 mM EDTA, 0.44 M KCl, 25% Glycerol), and centrifuged at 3000 rpm for 10 min. Supernatants were then dialyzed against Buffer D (20 mM Hepes, pH7.9, 100 mM KCl, 0.2 mM EDTA, 20% glycerol), and centrifuged at 3000 rpm for 10 min to remove particulate material.

**EMSA and UV-crosslinking immunoprecipitation:**

Oligonucleotide sequences for EMSA probes (lower case nucleotides indicate mutations) were as follows: AP1 sense 5’-GTGCCGGTGACTCACCAGGTTT-3’, AP1 antisense 5’-TTTAAACTCGTGAGTCACCGGC-3’, AP1m sense 5’-GTGCCGGTgAATGAGGTTT-3’, AP1m antisense 5’-TTTAAACTCaTGATcCACCGGC-3’, AP1 sense for crosslink 5’-GTGCCGGBrGACBrCACGAGGTTT, AP1m sense for crosslink 5’-GTGCCGGBrGgBrCAtGAGGTTT, CRE sense 5’-GCCCAGCTGACTGACGTCACCAAT-3’, CRE antisense 5’-TAGTATTGGTGACGTACCCAGT-3’, CREm sense 5’-GCCCAGCtGACGACTCACCAAT-3’, CREm antisense 5’-TAGTATTGGTGgaGTACCCAGT-3’

Briefly, annealed oligonucleotides were labeled with $^{32}$P-dATP using the DNA polymerase Klenow fragment. After incubation with nuclear extract for 15
min, the binding complexes were resolved on 6% native PAGE. For supershift assay, 4 ul α-CREB antibody was added to the reaction mixture for an additional 30 min before resolving on PAGE.

For AP1 cross-linking-immunoprecipitation, the reaction mixture was crosslinked using 310nm UV transilluminator (Fisher Scientific) for 30 min. The crosslinked complexes were then immunoprecipitated using α-c-Fos, α-c-Jun or α-β-actin antibody in non-denaturing RAF solution (20 mM Tris, 139 mM NaCl, 10% glycerol, 1% NP40, pH 8.0). The precipitated complexes were collected for scintillation counting.

**RNase Protection Assays:**

Total RNA was isolated 36-41 hours post-transfection as previously described (42, 57). RNase protections were performed as previously described (42, 57), using homologous anti-sense probes, termed “RP”, which spanned nts 1767-1906 for the AAV2 promoter and nts 1846-1985 for the P41 promoter of AAV5, respectively. The probes designated “DM” start and end at the same original nucleotide position but contain the corresponding AP1 and CRE debilitating mutation (for AAV5) or enabling mutation (for AAV2).
RESULTS

The core capsid promoter of AAV5 is stronger than that of AAV2.

Previous comparison indicated that in Ad5 E1A- and E1B-expressing 293 cells, the basal activity within RepCap plasmids of the AAV5 P41 capsid gene promoter was significantly greater than that of the P40 capsid gene promoters of AAV serotypes 1, 2, 3, 4 and 6 in similar backgrounds (76). To systematically identify cis-acting sequences that might account for these different expression levels, we used luciferase reporter assays to first compare the core promoter activities of AAV2 P40 and AAV5 P41 (Fig 1).

The major determinants of both core promoter activities (as diagrammed in Figure III-1A) were found, as expected, to be the TATA box and initiator element. Both elements were required for full promoter activity, as mutation of either element in either of the core promoter constructs CP40 or CP41 decreased activity to about half of the wild-type level (compare AV2 CP40 to AV2 CP40 TATAm and AV2 CP40 Inrm, and AV5 CP41 to AV5 CP41 TATAm and AV5 CP41 Inrm, Figure III-1B). Consistent with our previous evaluations, the AAV5 core promoter had about 5-times greater activity than the AAV2 core promoter (compare AV2 CP40 and AV5 CP41, Figure III-1B).

To further identify which elements within the core promoter regions might determine the difference in their activity, we exchanged their TATA boxes and initiator regions plus associated flanking regions. Replacing the AAV2 TATA box with the AAV5 TATA box resulted in greater than 3-fold higher activity compared to the wild-type AAV2 core promoter (Figure III-1B, compare AV2 CP40 with AV2
Figure III-1. Core promoter comparison of AAV2 P40 and AAV5 P41. (A)

Sequences of AAV2 P40 core promoter (CP40 nt 1823 to 1887) and AAV5 P41 core promoter (CP41 nt 1880 to 1970). The core promoters start from the TATA box of each promoter, and the luciferase genes from pGL3 are cloned downstream for the reporter assay below. TATA box and initiator sequences are indicated in boldface letters. The vertical lines in the boxes indicate the borders of TATA box and initiator elements in the hybrid promoters (AV2 CP40 V5TATA, AV2 CP40 V5Inr, AV5 CP41 V2TATA, and AV5 CP41 V2Inr) tested below. (B) Luciferase (Luc) activity of the wild-type, mutant, and hybrid promoters in 293 cells. The activity of each promoter is the average of duplicate samples from at least three experiments. All values are standardized to Renilla luciferase (R-Luc) activity generated by a TK promoter-driven Renilla luciferase expression vector used as an internal control.
(A) TATA Box  Initiator

AV2 CP40

\[
\text{TATAAGTGAGCCCAACGGTGCCCAGGAGTCAAGTTGCGAGCCCATCGACGAGCGGAAGCTT-} \text{Luc}
\]

1823  \ G \ G  \ 1887

AV5 CP41

\[
\text{TATAAATCTGGAGAGCTGGCCAGCAGTCATTGTCACGAGGCAGTTGAGTGATGCATGGAC-} \text{Luc}
\]

1823  \ G \ G  \ 1883  \ 1970

(B)

![Bar graph showing Luc/Rluc levels for different conditions](image)
CP40 V5TATA), whereas replacement of the AAV2 initiator with the AAV5 initiator resulted in a chimera with only about half of the activity of the wild-type (Figure III-1B, compare AV2 CP40 with AV2 CP40 V5Inr). In reciprocal experiments, replacement of the AAV5 TATA box with that of AAV2 resulted in a chimera with only half of the original activity (Figure III-1B, compare AV5 CP41 with AV5 CP41 V2TATA), while replacement of the AAV5 initiator sequence with that of AAV2 resulted in only a slight increase in activity (Figure III-1B, compare AV5 CP41 with AV5 CP41 V2Inr). These results indicated that AAV5 contains a stronger TATA box while AAV2 has a stronger initiator, yet the wild-type AAV5 core promoter had an overall stronger activity than the wild-type AAV2 core promoter. These differences, however, did not seem strong enough to explain the differences in expression observed in the parent RepCap constructs (76), and so further comparison of the upstream regions of the two promoters was performed.

**AP1 and CRE elements in AAV5 P41 upstream sequence contribute to its high basal activity.**

We next examined the roles of the upstream regions of AAV2 and AAV5 by linking various portions of these regions to AAV2 or AAV5 core promoters driving the luciferase gene. An alignment of the upstream regions of AAV5 P41 and AAV2 P40, designed to maximize sequence identity, is shown in Figure III-2A. These regions show significant differences, and identity between the two
Figure III-2. Upstream sequence analysis of the capsid promoters. (A)

Upstream sequence alignment of AAV2 P40 promoter (USP40 nt 1655 to 1822) and AAV5 P41 promoter (USP41 nt 1681 to 1879). Identical nucleotides are shaded, and the abbreviations L, S, and min indicate the different lengths of the upstream sequences linked to the core promoters in the luciferase constructs tested below. The consensus AP1 and CRE sites are also indicated in the boxes. (B and C) AAV5 upstream sequence enhances promoter activity greater than AAV2 upstream sequence. Different lengths of AAV2 and AAV5 upstream sequences are linked to core promoter CP40 (B) or CP41 (C), and the activity of each promoter is the average of duplicate samples from at least three experiments. All values are standardized to Renilla luciferase (R-Luc) activity generated by a TK promoter-driven Renilla luciferase expression vector used as an internal control. (D and E) AP1 and CRE are the major elements enhancing AAV5 P41 promoter activity. AAV2 core promoter CP40 is linked with wild-type AAV2, AAV5 upstream sequence, or AAV2 upstream sequence mutated to contain AP1, CRE, or both sites (D). A similar set was tested with AAV5 core promoter CP41 linked with wild-type AAV2, AAV5 upstream sequence, or AAV5 upstream sequence with AP1 and with the CRE site mutated (E).
drops dramatically downstream of nt 1744(AAV2)/1770(AAV5), the region
designated as “min” in Figure III-2A.

AAV2 or AAV5 upstream elements starting at either nt
1655(AAV2)/1681(AAV5) (L), 1708(AAV2)/1733(AAV5) (S), or
1744(AAV2)/1770(AAV5) (min), stimulated basal activity of the core promoters to
various extents (Figure III-2B, 2C, compare CP40 and CP41 to their relevant
chimeras), however, all AAV5 upstream sequences tested consistently increased
both AAV5 and AAV2 core promoter activity to a greater extent than AAV2
upstream sequences [compare 2L, 2S, and V2 min chimeras to 5L, 5S, V5min
chimeras, for CP40, Figure III-2B; and CP41, Figure III-2C). Thus, it was likely
that the AAV5 region from nt 1770 to the TATA box contained elements absent
from the analogous region in AAV2 that stimulated basal activity of these core
promoters.

Inspection of the region downstream of AAV5 nt 1770 identified consensus
binding sites for transcription factors AP1 and CREB (21) which are absent from
AAV2. To determine if these sites influence promoter activity, we carried out
mutational analysis of the AAV5 min and AAV2 min constructs described above.
The AAV5 AP1 site was destroyed by changing it to the resident AAV2 sequence
(VP5min AP1mCP41), and the CRE site was destroyed by changing the GT
nucleotides at nt 1865-66 to TC (V5minCREmCP41), a mutation shown by others
(54) to prevent CREB binding. The double knockout mutant was also made
(V5minAP1mCREmCP41). Reciprocal mutations were made in AAV2 min by
changing the region corresponding to the AAV5 AP1 site (nts 1761-66) so as to
add the AAV5 sequence (V2minAP1CP40), or by adding the AAV5 CRE site directly downstream of AAV2 nt 1822 (V2minCRECP40). The double addition mutation was also made (V2minAP1CRECP40).

As predicted, the AAV5 promoter with either AP1 or CRE mutated had reduced activity compared to the P41 core promoter containing the wild-type AAV5 upstream sequence (compare V5minAP1mCP41 and V5minCREmCP41 to V5minCP41, Fig 2E). The double knockout mutation decreased promoter activity even further, to approximately the level generated by AAV2min (compare V5minAP1mCREmCP41 to V2minCP41, Fig 2E). Within the AAV2 background, creation of either an AP1 or CRE site doubled the basal promoter activity (compare V2minAP1CP40 and V2minCRECP40 to V2minCP40, Fig 2D). When both sites were added, improvement almost to the level achieved by AAV5 min was observed (compare V2minAP1CRECP40 to V5minCP40, Fig 2D). These results suggested that the AP1 and CRE elements present in the AAV5 upstream region were at least partially responsible for the elevated basal activity of AAV5 P41 in 293 cells.

**The AAV5 AP1 and CRE elements bind their cognate factors in 293 cells.**

EMSA was used to investigate the binding of cellular factors to the AAV5 AP1- and CRE-binding sites. Both wild-type, but not mutant, probes were able to bind cellular factors present in 293-cell nuclear extracts (compare Figure III-3A and 3B, lanes 1 to 2). Excess cold wild-type, but not mutant, probes effectively
Figure III-3. Physical interaction between cellular factors and AP1 and CRE sites. (A and B) EMSA of AP1 and CRE with 293 cell nuclear extract. Free probes (F) and specific shifted (S) bands are indicated on the left. Complexes formed with wild-type and mutant (m) probes are shown in lanes 1 and 2, respectively. Tenfold cold wild-type (lane 3) and labeled mutant (lane 4) probe competition results are also included. The bands marked with an asterisk may represent additional unidentified complexes since they are removed by competition. (C) Supershift assay for CRE probe. In lane 2, after incubation of the probe and nuclear extract, 4 µl anti-CREB (a-CREB) antibody was added to the reaction mixture for an additional 30 min. A supershifted band is indicated by SS. (D) UV cross-link coupled immunoprecipitation to test the binding of c-Fos and c-Jun to the AP1 element. The radioactivity of the immunoprecipitated complexes was measured using scintillation counting. A wild-type AP1 probe with anti-β-actin (α-β-actin) antibody and an AP1 mutant probe (AP1m) with anti-c-Fos (α-c-Fos) or anti-c-Jun (α-c-Jun) antibodies were used as controls. The data shown are the average of duplicate samples from at least three experiments.
competed for binding (Fig 3A and 3B, compare lanes 3 and 4). To determine the specificity of these complexes, supershift assays were performed. AP1 sites most commonly bind to a heterodimer formed by c-Fos and c-Jun (25), while CREB is the primary factor that binds to CRE sites (15). A super-shifted complex was clearly detected for the CRE probe when α-CREB antibody was added (Figure III-3C lane 2), suggesting that CREB indeed forms a complex with this site in the AAV5 P41 promoter. α-c-Fos or α-c-Jun antibodies failed to generate detectable super-shifted bands complexing with the AAV5 AP1 element (data not shown), perhaps because of instability of the complex, however, UV-crosslinking of cellular complexes with this element allowed specific immunoprecipitation (compared to α-β-actin antibodies), by either α-c-Fos and α-c-Jun antibodies, of a wild-type radioactive probe to greater extents than a mutant radioactive probe (Figure III-3D), suggesting that c-Fos and c-Jun form a complex with the AAV5 AP1 site.

Activation by E1A is independent of the AAV5 AP1 and CRE sites.

As previously reported (76), the basal activity of AAV5 P41 is higher in 293 cells than HeLa cells. To test whether the adenovirus E1A and/or E1B gene products endogenously produced in 293 cells were responsible for this difference, the expression of P41 from a construct containing wild-type AAV5 upstream sequences (V5min CP41) was tested in HeLa cells, following co-transfection with various combinations of the Ad5 E1 genes. E1A, or E1A together with E1B, increased promoter activity in this assay, while E1B alone did

64
not, indicating that E1A was responsible for the activation effect (Figure III-4, left panel). The Ad E1A gene products interact with many cellular proteins (including CBP, which is a binding partner of CREB) to alter the transcription of both viral and cellular genes, although it does not bind DNA directly (20). Surprisingly, we found that while basal levels of expression were lower, a P41 reporter construct containing mutations of the AP1 and CRE sites (V5minAP1mCREm) was still activated to similar degrees by E1A (compare the first two column of each panel in Figure III-4). These results suggested that the increase in P41 expression in response to E1A was not mediated by the AP1 or CRE sites, but rather the effect of E1A was likely targeted directly to the basic transcription machinery.

The E1A gene generates 13S and 12S isoforms by alternative splicing which differ mainly in the retention of conserved region 3 (CR3) in the 13S isoform (20). Furthermore, the CR1 and N terminal regions of E1A, present in both isoforms, contain domains that interact with p300/CBP (73), which might be predicted to be important for CRE mediated E1A activation if such a pathway exists. We found that the 12S E1A product, supplied from a cDNA expression vector, increased promoter activity from each of the two reporter constructs tested to levels similar to that induced by the 13S E1A product (Figure III-4), suggesting that CR3 was not essential for its transactivating activity. Similarly, deletion of the CR1 region from the 13S cDNA expression vector did not abrogate its ability to activate either of the two reporter constructs. However, deletion of the N-terminal region from this protein did abolish its activating activity, although all mutant proteins were expressed at levels equivalent to the
Figure III-4. E1A activates AAV5 P41 promoter, but the activation is not directly mediated by AP1 or CRE sites. Shown is the luciferase activity of P41 reporter constructs with either the wild-type or AP1 and CRE mutant upstream regions, tested in HeLa cells together with cotransfection of different E1A isoforms, E1A mutants, or empty vector SK+. The activity is the average of duplicate samples from at least three experiments, as described in Materials and Methods.
wild-type protein (data not shown). These results suggested that the N-terminal domain was required for E1A function, consistent with a role for p300/CBP in this process.

The presence or absence of AP1 and CRE elements help determine expression of the capsid gene promoter within the viral RepCap background.

To further examine the roles of the AP1 and CRE elements, the debilitating mutations within AAV5 P41 and the enabling mutations within AAV2 P40 described above were introduced into AAV5 and AAV2 RepCap constructs. To prevent confounding expression of potentially mutant Rep proteins from these constructs, a premature termination codon was introduced into the NH2-terminal region of the large Rep proteins, as indicated (Figure III-5A), and activation of these constructs was then evaluated in the presence or absence of the AAV2 and AAV5 Rep proteins provided in trans. The AAV2 P5 promoter region provides a Rep binding element (RBE) competent for directing Rep activation, however, the AAV5 P7 region does not, and hence simple AAV5 RepCap constructs would not be expected to support Rep activation (76). The AAV5 ITR, however, does contain an RBE which effectively functions as a binding platform, and supports AAV5 Rep activation of weak promoters (76). Therefore, serotype-matched, RBE-containing half-TR sequences were added at the 3’ end of the four reporter constructs to provide RBEs for effective Rep activation, yet not
Figure III-5. AP1 and CRE enhance promoter activity in the viral RepCap background. (A) Illustration of the constructs used in the assay. AAV2 RepStopCapTR contains AAV2 sequence nt 145 to 4626 and a premature stop codon at nt 489. AAV2 RepStopDMCapTR has additional AP1 and CRE sites present in the P40 region. AAV5 RepStopCapTR contains AAV5 sequence nt 185 to 4587 and a premature stop codon at nt 480. AAV5 RepStopDMCapTR has additional AP1 and CRE mutation to knock out those sites in the P41 region. The probes used in the RNase protection assay are also shown. (B) RNase protection assay, performed as previously described (42, 57), using homologous RP probes to protect RNA generated in 293 cells following transfection of the constructs mentioned above. pBluescript SK+ was used as an empty vector control. AAV2 or AAV5 Rep was provided in trans (V2 or V5, respectively), driven by human immunodeficiency virus promoter, as previously described (51, 52). pEGFPC1 (Clontech, Mountain View, CA) was cotransfected as an internal control. Transfection of samples for lanes 2 to 4, 6 to 8, 10 to 12, and 14 to 16 was followed by adenovirus infection to allow activation, as indicated. (C) Quantification, using Fujifilm MultiGauge software, of RNase protection. A representative example is shown. Data from at least three experiments, with standard error bars, are presented as the activation (fold) of P40 (P41) promoters comparing transcription levels. The wild-type P40 (P41) level without Rep cotransfection and adenovirus infection is set as 1.
allow replication which would introduce another variable into the evaluation of expression levels. Expression of these constructs was then tested in 293 cells in the presence and absence of co-transfected HIV promoter-driven Rep-expressing plasmids (51, 52).

Similar to previous results (51, 76), we observed a much higher basal level of wild-type AAV5 P41 than wild-type AAV2 P40 from these constructs in 293 cells (Figure III-5B, compare lane 1 to lane 9), and the basal activity of both wild-type promoters was stimulated only slightly by adenovirus (Figure III-5B, compare lanes 2 to 1; and 10 to 9). The addition of AAV5 Rep slightly increased wild-type P41 activity (about 3 fold) (Figure III-5B, compare lane 4 to lane 2), while added AAV2 Rep had a negligible effect on P41 (Figure III-5B, compare lane 3 to 2). Mutation of the AAV5 AP1 and CRE binding sites significantly decreased P41 basal activity both in the absence (to approximately 1/10 of the original level, Fig 5B, compare lane 5 to lane 1), and presence (to approximately 1/17 of the original level, Figure III-5B, compare lane 6 to lane 2) of adenovirus. Interestingly, this weakened promoter was activated by AAV5 Rep (more than 7 fold), but not by AAV2 Rep (Figure III-5B, compare lanes 8 and 7 to lane 6). These results demonstrate the importance of the AP1 and CRE elements in maintaining high basal promoter activity in 293 cells. In addition, these results confirm that AAV5 Rep can act as a transcriptional activator under certain circumstances (a weaker promoter and a good binding site) as previously described (76).
The addition of an AP1 and CRE binding site conferred to the AAV2 P40 promoter a much higher basal activity, both in the absence and presence of adenovirus (Figure III-5B, compare lane 13 to lane 9 and lane 14 to lane 10). Both wild-type and AP1 plus CRE site-containing promoters were activated by AAV2, but not AAV5, Rep (Figure III-5B, compare lanes 11 and 12 to lane 10; lanes 15 and 16 to lane 14). While the final activated levels of the AP1 and CRE binding site-containing P40 promoters was higher than for the wild-type P40, the activation fold induced by AAV2 Rep was somewhat lower for the AP1 plus CRE site-containing P40 compared to the wild-type (3.5 fold vs 5 fold, Figure III-5B, compare lane 11 to 10 and lane 15 to 14). This result suggested that the addition of AP1 and CRE sites did not disrupt the elements required for Rep activation; however, activation levels were lower due to increased basal activity.
DISCUSSION

The capsid promoter of AAV5 differs from that of the AAV2–like Dependoviruses (AAV1, 3, 4 and 6) in that it exhibits high constitutive activity in Ad5 E1A and E1B-expressing 293 cells, but not HeLa cells, either as part of a minimal P41-expression construct, or a larger, partial ITR-containing, RepCap construct. We show in this manuscript that this is due to both a stronger core promoter and the presence in AAV5 P41 of CREB and AP1 binding sites immediately upstream of the core sequences. Increased expression from the core elements depends upon the Ad5 E1A protein, but activation due to the upstream CREB and AP1 binding sites is independent of E1A.

The AP1 and CREB binding sites are found upstream of the capsid gene promoters of other AAV5-like Dependoviruses, including bovine AAV (B-AAV), and caprine AAV (Go.1-AAV), (Fig 6). These elements are lacking from all of the AAV2-like viruses, as well as the avian AAV (A-AAV), which is the most divergent of the non-primate AAV5 like viruses (49). It has become clear that the AAV5-like AAVs (including those non-primate AAVs so far characterized) are quite distinct from the AAV2-like viruses, both phylogenetically, and in genetic organization (48, 49). The presence of the AP1 and CREB sites, and hence differences in capsid promoter activity, is another feature that discriminates between these two groups, and which potentially can be used to evaluate evolutionary distances. Although the animal AAVs have been shown to co-circulate with adenovirus species (48, 49), the natural host for AAV5 has not
Figure III-6. Genome sequence comparison of AAV2 and AAV5-like viruses.

The P40/P41 upstream sequence alignments of AAV2, AAV5, bovine AAV (BAAV), and caprine AAV (Go-AAV) are shown. The upstream promoter sequences from each virus were aligned using Vector NTI software (Invitrogen) with AAV5 as the consensus reference. Identical nucleotides in all AAVs are shown in the darkest shade of gray. Nucleotides identical to the AAV5 sequence are shown in the lighter shade of gray. The consensus AP1 and CRE sites are indicated in boxes.
been established. It remains possible that the response of AAV5 to HSV may be different.

AP1 and CREB have been extensively studied as part of cellular signal transduction pathways in many different systems (15, 25). Both sites have been found in the adenovirus E3 promoter and have been implicated in the stabilization of TATA box binding complex (54). Because of the intimacy between the AAV and Ad life cycles, it is perhaps not surprising to find that they perform at least in part a similar function during AAV5 gene regulation. As found previously for the Ad5 E3 promoter (54), the disruption of the AAV5 AP1 and CREB binding sites did not eliminate activation by E1A, which suggested that E1A may affect the general transcription machinery directly at the core promoter, and thus synergistically enhance promoter activity with transcription factors bound upstream.

Regulation of transcription from an incoming infectious viral genome may ultimately show differences compared to the plasmid transfection results shown here. However, in AAV5 RepCap constructs bearing a single RBE-containing partial AAV5 ITR sequence, the presence of CREB and AP1 binding sites upstream of the capsid gene promoter is the main determinant to its high basal activity. Furthermore, AAV5 Rep only slightly increases P41 activity in this background in the presence of Ad5. When the AP1 and CREB binding sites in this background were disrupted, the basal level of expression of P41 was severely decreased, and interestingly, in the mutant background AAV5 Rep, but not AAV2 Rep, was able to significantly activate this promoter although both Rep
proteins have been shown to bind to the RBE resident on the AAV5 ITR with similar affinities (10). In a reciprocal set of experiments, introduction of these transcription factor binding sites into the AAV2 background containing the AAV2 half ITR increased the basal activity of the capsid promoter thus decreasing the relative activation fold in response to AAV2 Rep. AAV5 Rep remained a very poor activator of P40 in this context. Although the AAV2 and AAV5 Rep proteins have been shown to bind both the AAV5 and AAV2 ITRs with similar affinities (10), activation by Rep could certainly be influenced by a specific configuration adopted during the binding of each Rep to its own ITR (26, 72). The two Rep proteins have conserved their activation ability, however, the magnitude of activation seems to be determined both by the relative strength of the target promoters and the nature of its cognate binding site.

It may be that relative expression level of the capsid gene promoter of AAV5 needs to be more significantly regulated during viral infection, and thus capsid gene expression during infection by wild-type viral genomes that contain two ITRs may be different. In fact, following transfection of the AAV5 infectious clone, the basal activity of P41 was found to be relatively lower, and greater (though still modest) levels of transactivation was observed from these full length molecules (51). However, it should be noted that while the genomes of both MVM and AAV2 are expressed in a temporal manner, i.e., the capsid encoding genes are upregulated by the prior expression of the nonstructural proteins, steady-state levels of promoter activity is reached quite quickly, after which there
is little change in the relative strength of the viral promoters or the ratio of the
abundance of the various transcripts.
IV. BOVINE ADENO-ASSOCIATED VIRUS EXHIBITS DISTINCT REGULATION OF CAPSID GENE PROMOTER DESPITE CLOSE RELATIONSHIP TO ADENO ASSOCIATED VIRUS 5

INTRODUCTION

As members of *Dependovirus* genus of the *Parvoviridae* family, adeno-associated viruses (AAV) depend on helper viruses such as adenovirus or herpes simplex virus (HSV) to complete their life cycle. Different AAVs utilize diverse strategies to modulate their capsid promoter activity. The AAV2 capsid promoter P40, which has been extensively studied and similar to many human AAVs, has a low basal transcription activity and is activated by its own non-structural Rep protein via a binding platform provided by its inverted terminal repeats (ITR) or P5 promoter in the presence of helper viral functions (30, 46). The AAV5 capsid promoter P41, on the other hand, has a higher basal activity in E1A/B expressing 293 cells, partly due to the presence of AP1 and CREB transcription factor binding sites upstream of the promoter and is not further activated by its Rep (75).

AAV5, the most divergent serotype of all human AAVs, is closely related to animal AAVs, such as caprine AAV (Go.1-AAV) and bovine AAV (BAAV), which were originally isolated from animal adenovirus stocks (1, 14, 33, 41). The nucleotide sequence of AAV5 and BAAV is more than 77% identical, which is greater than the homology between AAV5 and any other serotypes of human AAV (AAV1, 2, 3, 4 and 6). Examination of the P41 capsid promoter region revealed the presence of both AP1 and CREB binding sites in BAAV (75), yet in
contrast to AAV5, it is still activated by BAAV Rep targeted to the transcription template in the presence of adenovirus (49). This observation was unexpected given the high homology between AAV5 and BAAV sequences. Therefore, in this study, we systematically compared the transcription profiles of the two viruses under the same conditions. The activation of BAAV P41 promoter by its own Rep and interestingly, by AAV5 Rep as well, was confirmed. Although mutation of both the BAAV P41 AP1 and CRE decreased its activity, the lower basal level of BAAV P41 was determined to be due to a combination of an intrinsic weakness in the promoter itself and possible effects of surrounding sequences. In addition, we observed elevated upstream transcription in constructs containing an RBE at its left-hand end [a left-hand (1/2) ITR], which for AAV5 was due to increased transcription of the P7 promoter, and for BAAV, due to a combination of P7 and P19 transcription in BAAV. These results revealed significant difference in transcription control between these two viruses despite their high sequence homology.
MATERIALS AND METHODS

Cells and Virus:

293 cells were propagated as previously described (40). Transfections, using Lipofectamine and the Plus reagent (Invitrogen, Carlsbad, CA), were performed as previously described (52), and when Ad5 was co-infected, this was done 5 hours after transfection at a moi of 5.

Plasmid Constructs:

RepCap and RepStopCap plasmids: AAV5 RepCap (containing AAV5 nts 185-4448) has been described previously (51). BAAV RepCap (containing BAAV nts 193-4493) was constructed by inserting BAAV sequence between SacII-KpnI sites on pBluescript SK(+) (Stratagene, La Jolla, Calif.). For AAV5, TRRepCap extend AAV5 RepCap at the 5’ end to nt 75, RepCapTR extend AAV5 RepCap at the 3’ to nt 4591 and TRRepCapTR extends both ends. For BAAV, TRRepCap extends the 5’ end to nt 81, RepCapTR extends the 3’ end to nt 4637 and TRRepCapTR extends both ends as well. Stop codons were introduced at nt 480 in the AAV5 constructs and nt 54 in the BAAV constructs mentioned above to make their RepStopCap counterparts. AAV5 and BAAV Rep expression plasmids were mentioned before (49, 51), which contains silent mutation in the RP region to avoid hybridization with the aforementioned test plasmids. SacII was used to linearize the AAV5 and BAAV plasmids, except TRBAAVRepCapTR plasmid, which was linearized by KpnI.

Luciferase reporter plasmids: All luciferase reporter constructs are based on pGL3-Basic (Promega, Madison, WI). All AAV5 reporter constructs were
mentioned previously (75). BAAV reporter constructs utilize the same cloning strategy but use the BAAV counterpart. BAAV core promoter (CP41) contains nt 1889-1979. BAVmin added BAV upstream sequence (nt 1779-1888) upstream of the BAAV CP41 construct. BAVmin AP1m is based on BAVmin with the sequence ACTCAC at nt 1797 to 1802 changed to GAGCAT. BAVmin CREm is based on BAVmin with the GT at nt 1875 to 1876 changed to TC. BAVmin AP1m CREm contains both changes.

AAV5-BAAV chimeric plasmids: chimeras were based on AAV5 and BAAV RepStopCap plasmids. The border (AAV5 nt 1472 and BAAV nt 1481) was chosen at a significant homologous region of the two constructs. BAAVRepStopCapAAV5P41 was based on BAAVRepStopCap background with AAV5 P41 sequence (nt 1472-1974) replacing BAAV P41 sequence (nt 1481-1983). AAV5RepStopCapBAAVP41 was based on BAAVRepStopCap background with the reverse exchange.

RNase Protection Assays:

Total RNA was isolated 36-41 hours post-transfection as previously described (9,16). RNase protections were performed as previously described (9,16), using homologous anti-sense probes. AAV5 P7 probe spans nt 278-431, P19 probe spans nt 839-996 and RP probe spans nt 1846-1986. BAAV P7 probe spans nt 280-440, P19 probe spans nt 848-1005 and RP probe spans nt 1855-1995.
**Luciferase Assays:**

Luciferase assays were performed according to the manufacturer's suggested protocols (Promega, Madison, WI). Briefly, 293 cells grown in 12-well plates were transfected with 0.1 µg per well of the luciferase reporter constructs along with 0.35 µg per well of empty pBluescript SK(+) (Stratagene, La Jolla, CA) using the Lipofectamine reagent (Invitrogen). In addition, 0.05 µg per well of thymidine kinase (TK)-driven *Renilla* luciferase gene reporter was cotransfected as an internal control. The total amount of DNA transfected into each well was kept at 0.5 µg. Thirty-six hours after transfection, cells were lysed and the luciferase activity was tested using the Promega dual-luciferase reporter assay system (Promega). Each experiment represents the average of duplicates from three individual experiments (error bars shown), each normalized to *Renilla* luciferase activity.
RESULTS

*Basic observation of the transcription profiles of BAAV compared to AAV5:*

Previously we have shown that in E1A- and E1B-expressing 293 cells, ITR-deleted non-infectious AAV5 RepCap plasmid exhibited a high basal level of transcription from its capsid promoter P41 which, unlike AAV2 P40, was not further increased by helper adenovirus infection (76). For similar BAAV constructs, however, its P41 promoter was only slightly enhanced by its Rep protein (49). When an RBE was inserted in this construct in the form of [1/2] ITR sequence, activity of BAAV P41 was significantly increased by its own Rep, presumably because this sequence provided a binding platform to target Rep to the transcription template for activation. This was unexpected because the BAAV nucleotide sequence is more homologous to AAV5 than to other human AAVs, yet this activation pattern was more similar to that of AAV2. Therefore, we decided to comprehensively compare the regulation of the BAAV and AAV5 P41 promoters by placing them in a similar background, as diagrammed in Fig 1A. In addition to characterizing the ITR-deleted RepCap constructs previously described (49, 51), we also extended the constructs at the 5’ and/or 3’ end so that the RBE-containing [1/2] ITR sequence was added to provide Rep binding platform yet still lacks essential cis-elements required for genome replication. For both sets of plasmids, P41 transcription levels in the presence or absence of adenovirus infection were compared using RNase protection assays and the quantification of these results are shown in the bar graph in Fig 1C and E.
**Figure IV-1. Basic transcription profile comparison between BAAV and AAV5 in response to adenovirus infection.**

(A) Diagrams of RepCap constructs used in this experiment. P7, P19 and P41 promoters are indicated by arrowheads. Probes across P41 region used in the RNase protection assay are labeled as RP (spanning nt 1855 to 1995 of BAAV and 1846 to 1986 of AAV5). [1/2]ITR sequences are placed at the left (nt 81 to 192 of BAAV and 75 to 184 of AAV5) or right hand (nt 4494 to 4637 of BAAV and nt 4449 to 4591 of AAV5) of the genome to allow Rep binding. (B) and (D) RNase protection assay, performed as previously described (9,16), using aforementioned RP probe across P41 initiation site of RNA generated in 293 cells following transfection of BAAV(B) and AAV5(D) constructs diagramed in (A). pEGFPC1 (Clontech, Mountain View, CA) was cotransfected as internal control. Transcription signals from GFP, upstream promoters (ITRP7P19) and P41 are labeled. (C) and (E) Quantification, using Fujifilm MultiGauge software, of RNase protections, a representative of which is shown. Data from at least three experiments, with standard error bars, are presented as the level of P41 transcripts normalized to GFP signal.
For BAAV, P41 transcription from RepCap plasmids was slightly increased by adenovirus infection (Fig 1B, compare lane 2 to 1). Consistent with previous observations (49), more significant activation by adenovirus was observed in constructs with the [1/2] ITR sequence present, regardless its position (Fig 1B, compare lane 4, 6, 8 to lane 3, 5, 7, respectively). P41 transcription from AAV5, however, was not significantly increased by adenovirus infection in any of the constructs (Fig 1D, compare lane 2, 4, 6, 8 to lane 1, 3, 5, 7, respectively). Therefore, although genetically closer to AAV5, the transcription pattern of BAAV was more similar to that of AAV2, whose large Rep proteins (Rep 78 and 68) binds a Rep binding element (RBE) and recruits transcription factors for capsid promoter activation (76). Based on our previous comparison study between AAV2 and AAV5, this may have been due to the difference of activation ability of the Rep proteins (76) and/or cis sequences around P41 promoters (75). We also noticed that for both BAAV and AAV5, the upstream transcripts from constructs containing left-hand [1/2] ITR sequences (TRRepCap and TRRepCapTR) were significantly more abundant than those from RepCap and RepCapTR. This will be further investigated in the latter part of the study.

**BAAV and AAV5 Rep have similar activation capabilities:**

To compare the potential role of the BAAV and AAV5 Rep proteins in P41 activation, we engineered a set of RepStopCap constructs, based on the RepCap constructs mentioned above, by introduction of premature termination codon downstream of the P7 promoters (Fig 2A). To test the role of AAV5 and
Figure IV-2. Transcription activation by AAV5 and BAAV Rep protein. (A)

Diagrams of RepStopCap constructs and probes used in this experiment. Constructs are the same as in Figure 1A except the introduction of premature termination codon at nt 541 for BAAV (49) and nt 480 for AAV5 (76). (B) and (D) RNase protection assay of the constructs cotransfected with SK empty-vector control or AAV5 or BAAV Rep expression constructs as activators, followed by adenovirus infection. pEGFPC1 is cotransfected as control. (C) and (E) Quantification of the protection shown above. Data from at least three experiments, with standard error bars, are presented as the level of P41 transcripts normalized to GFP signal.
BAAV Rep proteins as the activators, they were provided in trans with adenovirus infection. As expected, all BAAV RepStopCap constructs, regardless of the placement of the [1/2] ITR sequence, produced very low level of P41 transcripts in the absence of Rep proteins (Fig 2B, lane 1, 4, 7, 10). Expression of P41 in RepStopCap was not increased by AAV5 or BAAV Rep due to the lack of Rep binding sequences in this construct (Fig 2B, compare lane 2 and 3 to lane 1). However, in constructs containing [1/2] ITR sequences competent for Rep binding (TRRepStopCap, RepStopCapTR and TRRepStopCapTR), P41 transcription was significantly increased by cotransfection of either AAV5 or BAAV Rep (Fig 2B, compare lane 5, 6 to lane 4, lane 8, 9 to lane 7 and lane 11, 12 to lane 10, respectively). These results confirmed the role of the Rep protein as an activator for the BAAV P41 promoter, and the similar activation fold obtained suggested that AAV5 and BAAV Rep have similar activation capabilities.

A similar set of experiments were carried out in the AAV5 background. Interestingly, we see a similar activation pattern as BAAV. P41 promoters in RepStopCap constructs were not activated by AAV5 or BAAV Rep (Fig 2D, compare lane 2 and 3 to lane 1), but P41 promoters in constructs containing [1/2] ITR sequences were activated by both AAV5 and BAAV Rep to a similar level (Fig 2D, compare lane 5, 6 to lane 4, lane 8, 9 to lane 7 and lane 11, 12 to lane 10). This was surprising as we did not observe P41 activation by adenovirus in the previous experiment. One possibility, therefore, was that AAV5 Rep activation of P41 promoter was independent of adenovirus. Thus, the same set of
experiments was repeated without adenovirus infection. In this instance, we did observe similar activation patterns of P41 by AAV5 and BAAV Rep in constructs with [1/2] ITR sequences present (data not shown), although the overall activation fold was slightly lower. Therefore, this activation by Rep requires Rep binding sequence but is likely to be independent of adenovirus infection.

**Cis sequences that contribute to the transcription regulation differences of BAAV and AAV5:**

As we have seen in Fig 1, BAAV and AAV5 respond differently to adenovirus infection. From previous comparison studies between AAV2 and AAV5 capsid promoters, we learned that the lower basal level of AAV2 P40 promoter allows for a high activation fold while the natively stronger AAV5 P41 promoter is not activated as much by adenovirus. Disruption of the AP1 and CRE sites in the AAV5 P41 upstream region debilitates basal promoter activity, yet allows a higher activation fold by AAV5 Rep (76). Therefore, we decided to compare BAAV and AAV5 P41 basal transcription in a similar setting. Sequence alignment of the BAAV and AAV5 P41 promoter upstream sequences (Fig 3A) revealed high homology between the two, with AP1 and CRE sites present in both. Luciferase assays showed that the upstream sequence functioned similarly in BAAV as in AAV5, which significantly increased promoter activity over the core promoter sequences alone (Fig 3B, compare minCP41 to CP41); although overall promoter activity of BAAV P41 was lower than AAV5 P41. Mutation in the AP1 or CRE sites decreased the promoter activity to about half of the wild-type


Figure IV-3. Analysis of cis sequences of BAAV and AAV5 that control P41 transcription. (A) Alignment of BAAV (nt 1779 to 1888) and AAV5 (nt 1770 to nt 1879) P41 promoter region. Consensus nucleotides are shaded. AP1 and CRE elements previously reported to mediate constitutive high level of AAV5 P41 promoter are also indicated. (B) Mutation analysis of P41 promoter. Luciferase gene driven by core promoter (CP41) starting with TATA box, or core promoters with minimum wild type or AP1/CRE mutant upstream sequences were transfected into 293 cells. Luciferase activity is normalized to Renilla luciferase (R-luc) transfection control. The values shown are the average with standard deviations from three independent plasmid transfections done in a single experiment. (C) Diagrams of wild type and chimeric RepCap constructs used in the assay below. Blank bar stands for AAV5 sequence while black bar stands for BAAV. All constructs contain premature termination codon downstream of P7 promoter (nt 541 for BAAV and nt 480 for AAV5). RP probe across P41 region is used for RNase protection assay. (D) RNase protection assay of the constructs cotransfected with SK empty-vector control or AAV5 or BAAV Rep expression constructs as activators, followed by adenovirus infection. Transcription signals from GFP, P7, P19 and P41 are indicated. (E) Quantification of the protection shown above. Data from at least three experiments, with standard error bars, are presented as the level of P41 transcripts normalized to GFP signal.
level (Fig 3B, compare minAP1mCP41, minCREmCP41 to minCP41), and the
double mutant promoter had an even lower activity (Fig 3B, compare
minAP1mCREmCP41 to minAP1mCP41 and minCREmCP41), consistent with
what we observed previously for the AAV5 P41 promoter. Therefore, both the
AP1 and CRE sites were functional in maintaining basal transcription of the
BAAV P41 promoter. However, the trivial difference in activity between the BAAV
and AAV5 P41 promoter observed in this assay (less than 2 fold) suggested that
sequences other than the promoter region might also contribute the different
transcription profiles.

To address this possibility, we engineered two chimeric constructs in
addition to the wild-type BAAV and AAV5. The BAAVRepStopCapAAV5P41
contains AAV5 P41 promoter in the BAAV background while
AAV5RepStopCapBAAVP41 has BAAV P41 promoter in the AAV5 background
(Fig 3C). Although the chimeric junction was chosen in a homologous region
between the two to avoid potential disruption of Rep open reading frame, there is
still the possibility of incorrect folding of the chimeric Rep protein. Therefore, a
premature termination codon was introduced downstream of the P7 promoter
and the Rep proteins were provided in trans.

The basal level of BAAV and AAV5 P41 was significantly different. AAV5
P41 transcription was approximately 7 fold higher than BAAV (Fig 3D, compare
lane 7 to lane 1). This effect is not due to Rep because adding AAV5 or BAAV
Rep in trans did not change the level of transcription (Fig 3D, compare lane 2, 3
to lane 1, and lane 8, 9 to lane 7). When AAV5 P41 was placed in the BAAV
background, however, transcription was decreased to about half of the original level (Fig 3D, compare lane 4 to lane 1), and Rep provided in trans did not recover activity to parental levels (Fig 3D, compare lane 5, 6 to lane 4). However, the level was still higher than the wild-type BAAV construct due to the innately stronger basal AAV5 P41 transcription, examined in Fig 3B. The reciprocal experiment in which the BAAV P41 promoter was placed in the AAV5 background increased BAAV P41 activity to about twice that of the original level (Fig 3D, compare lane 10 to 7) although not quite to the level of wild-type AAV5 P41, likely due to the same reason mentioned above. Addition of Rep still had no effect (Fig 3D, compare lane 11, 12 to lane 10). Therefore, the differences between the basal promoter activities we observed are due to a combination of intrinsic promoter properties and the effect of surrounding sequences.

**The origin of increased upstream transcripts in constructs containing the left-hand [1/2] ITR**

As observed in Fig 1, both BAAV and AAV5 constructs containing left-hand [1/2] ITR sequences have elevated levels of upstream transcription. To examine the origin of these transcripts, we designed separate probes across the P7 and P19 promoters to differentiate native transcripts from transcripts reading around the circular parental plasmid template (ITR vs. P7 using P7 probes, ITR+P7 vs. P19 using P19 probes, Fig 4A).

BAAV P7 probe indicated that there were more abundant transcripts from upstream than from P7 promoter itself, although these transcripts were seen at
Figure IV-4. Upstream transcription analysis of BAAV and AAV5. (A)

Diagrams of RepCap constructs and probes used. The RepCap constructs are the same as in Figure 1A. P7 probe spans from nt 280 to nt 440 of BAAV and nt 278 to nt 431 of AAV5. P19 probe spans from nt 848 to nt 1005 of BAAV and nt 839 to nt 996 of AAV5. (B) and (D) RNase protection assay of BAAV or AAV5 constructs using P7 or P19 probes. GFP, upstream, P7 and P19 transcripts are labeled. (*) in (D) represents undigested AAV5 P7 probe. (C) and (E) Quantification of the protection assay. Data from at least three experiments, with standard error bars, are presented as the level of transcripts normalized to GFP signal.
similar levels regardless of the placement of the [1/2] ITR. These could have been read-through by-products of other promoters or transcripts derived from the [1/2] ITR sequences in constructs TRRepCap and TRRepCapTR. In these two constructs, addition of the left-hand [1/2] ITR did slightly increase P7 transcription (Fig 4B left panel, compare lane 3, 4 to lane 1, 2, lane 7, 8 to lane 5, 6). This slight increase was even less obvious when the P19 probe was used (Fig 4B right panel), since the upstream transcripts examined here combined both of the transcripts studied in the previous panel. However, we did notice a more pronounced difference in accumulation of the P19 transcripts. Constructs with left-hand [1/2] ITR sequences produced significantly more P19 transcripts than those without the left-hand [1/2] ITR (Fig 4B right panel, compare lane 11, 12 to lane 9, 10, lane 15, 16 to lane 13, 14).

The same set of experiments carried out in AAV5 background revealed surprisingly different results. The constructs with left-hand [1/2] ITR sequences had significantly higher levels of upstream transcripts (Fig 4D, compare lane 3, 4 to lane 1, 2, lane 7, 8 to 5, 6). Protection with the discriminating probes indicated that these mostly likely reflected read-through transcription of the circular plasmid in addition to the [1/2] ITR transcription. Similar to BAAV, overall P7 transcription was also weaker than upstream signals in AAV5, although constructs with left-hand [1/2] ITR produced more P7 transcripts (Fig 4D left panel, compare lane 3, 4 to lane 1, 2, lane 7, 8 to lane 5, 6). This was further confirmed by the P19 probe as well (Fig 4D right panel). Interestingly, P19 promoters in all the AAV5 constructs produced significantly fewer transcripts than they did within BAAV,
and there was no significant difference between constructs that contain a left-hand [1/2] ITR sequence and those that do not.

To eliminate the possible artifact of read-through transcription, further analysis was carried out using linearized plasmids (Fig 5). P41 transcription from both BAAV and AAV5 followed the same pattern as the circular plasmids; BAAV P41 in constructs with the [1/2] ITR binding site at the left-hand end were activated in the presence of adenovirus (Fig 5A, compare lane 20 to 19, 22 to 21 and 24 to 23), while AAV5 P41 was not activated in any of the constructs in the same setting (Fig 5B, compare lane 20 to 19, 22 to 21 and 24 to 23). However, upstream transcripts exhibited different features. BAAV transcription from the [1/2] ITR, P7 and P19 was minimal except for constructs with left hand [1/2] ITR sequences in the presence of adenovirus (Fig 5A, lane 4, 8, 12 and 16), which suggested activation of these element with helper viral functions. AAV5, on the other hand, mostly resembled the transcription pattern of the circular plasmids. There was insignificant P19 transcription, and increased P7 transcription in constructs with left-hand [1/2] ITR sequences (Fig 5B, compare lane3, 4 to lane 1, 2 and lane 7, 8 to lane 5, 6). However, the linear plasmids did not produce detectable transcripts initiating upstream of P7 promoter, indicating that those observed in previous experiment were indeed read-through products and that authentic transcription from the [1/2] ITR was below the limit of detection, which agrees with our previous study using luciferase reporter assay (49). These results suggested BAAV transcription from the [1/2] ITR and all three promoters
Figure IV-5. Transcription analysis of linearized BAAV and AAV5 plasmids.

(A) and (B) RNase protection assay of linearized BAAV and AAV5 construct transfection using P7, P19 and RP probes. GFP probes are used as internal control. (*) in (B) represents undigested P7 probe. (C) and (D) Quantification of the protection assay. Data from at least three experiments, with standard error bars, are presented as the level of transcripts normalized to GFP signal.
responds to adenovirus infection similar to AAV2, while AAV5 exhibits unique patterns despite greater sequence homology with BAAV.
DISCUSSION

The comprehensive comparison between AAV5 and BAAV in this study revealed various differences in transcription control in spite of their sequence homology. BAAV demonstrated a transcription activation pattern similar to that of AAV2 rather than the phylogenetically-closer serotype AAV5. However, given the fact that our experiments were all carried out in human 293 cells with human adenovirus infection, much is yet to be investigated about BAAV in its natural host cell line (such as MDBK, bovine kidney cells) with its natural helper virus, bovine adenovirus. Unfortunately, this approach is currently not feasible because available bovine adenovirus stocks contain high level of BAAV (49).

Among all the serotypes of AAVs investigated so far, AAV5 is most unique due to its close resemblance to animal AAVs and its independence of adenovirus for its transcription control. Although regulation under other helper viral infection such as HSV might be different, we have enough evidence to suggest that adenoviral gene products (E4orf6 and VA) affect AAV5 gene expression mainly at the post-transcription level (43). Compared to transcriptional control, this approach is less efficient as the products have to be made first and then regulated. However, this might reflect a transient stage in viral host migration from animal to human beings, while the more tissue culture adapted AAV2 system may have evolved to modulate its gene expression more efficiently. It would be interesting to examine whether the transcription profile of AAV5 alters through prolonged passage in cell culture.
V. ONGOING PROJECTS:

i. Inhibition of MVM Replication by APOBEC3A:

INTRODUCTION

The autonomous minute virus of mice (MVM) is an important member of the parvovirus family which causes growth retardation and death of neonatal and fetal mouse (62). Its replication does not need an additional helper virus, but does rely on the host cellular machinery to complete its cycle, which is exemplified by the requirement of progression into S-phase of host cells for the viral lytic cycle (56). The MVM genome is organized into two overlapping transcription units, with two promoters P4 and P38 driving the non-structural NS1/2 protein and structural VP1/2 protein expression, respectively (Fig V-1A).

MVM replication is governed by a rolling hairpin mechanism, which depends on host enzymes and cellular functions during the S phase of the cell cycle. Replication initiates rightward from the 3’ terminal hydroxyl of the hairpin structure, which serves as a primer to extend a complementary copy of the viral genome until it reaches the 5’ terminus, resulting in a covalently closed DNA replication form (RF). Then, the non-structural protein NS1 binds and nicks the strand at the terminal resolution site. The rearranged hairpin serves as primer for strand displacement synthesis, leading to the formation of a dimer duplex intermediate (dRF). Viral capsid protein then packages single stranded DNA (ssDNA) to form the virion (Fig V-1B). A previous report has shown that for an MVM mutant (pMVM(VP1/23451)) which does not produce VP1 or VP2 capsid proteins, viral replication is halted at the ssDNA production step even dRF is still
Figure V-1. (A) Genome organization of minute virus of mouse (MVM). P4 and P38 promoters are shown in arrowheads. NS1 and NS2 are produced from P4 transcripts, while P38 drives VP capsid protein production. (B) Replication of MVM. Diagrams of different replication intermediates are shown and labeled on the right.
(A) [Diagram of viral protein structures and sequences]

(B) [Diagram of viral RNA replication cycle]

106
made (68). This mutant is very useful to delineate the different stages of viral replication.

APOBEC family proteins deaminate cytidine in RNA and/or DNA. Human APOBEC family includes APOBEC1 (hA1), APOBEC (hA2) and APOBEC3A, B, C, D/E, F, G and H (hA3A-H), while mouse encodes APOBEC1 (mA1), APOBEC2 (mA2), and a single APOBEC3 (mA3). hA3F and hA3G inhibit HIV (Human immunodeficiency virus) infection by introducing hypermutation into the viral sequence and this effect is suppressed by the viral Vif protein (23, 59), which targets APOBEC3G for ubiquitination and degradation (31, 34, 37, 60, 66). APOBEC3 proteins have also been shown to inhibit retrotransposon(4) and ssDNA virus, such as AAV replication (9), although the mechanism is still unknown. One previous study has shown that specific packaging of species specific APOBEC3 protein contribute to their inhibitory specificity(19), etc. the Mason-Pfizer monkey virus Gag binds and packages mA3 but not rhesus A3G, thus its replication is only inhibited by mA3. One the contrary, the murine leukemia virus packages rhesus A3G into the virion, but not mA3, thus its replication is inhibited only by rhesus A3G.

Because of the similarity between MVM and AAV, it is possible that APOBEC may also inhibit MVM replication. This part of the study tests this hypothesis and tries to delineate the steps in viral life cycle that are directly targeted by APOBEC.
MATERIALS AND METHODS

Cells and transfection:

Murine A9 cells were propagated as previously described (39).

Transfection of A9 cells was as previously described using calcium phosphate method (39).

Plasmid constructs:

The MVMp parent plasmid used in this study was the original infectious clone derived by Merchlinsky et al. (38) and Gardiner and Tattersall (22). The pMVM(VP1/23451) was constructed previously which contains a deletion within the viral capsid region to disrupt both VP1 and VP2 production.

APOBEC3A and C106S constructs are kindly provided by Matt Weitzman from Salk Institute. Both are constructed in the pcDNA3 (Invitrogen, CA) background.

Southern and Western analysis and RNase protection assays:

Southern analysis and Western blot analysis were performed exactly as previously described (39, 57). The complete homologous genomic clones were used as probes for Southern analysis, and the antibodies for NS1/NS2 Western analysis was polyclonal rabbit antibodies raised to NS1 and NS2, respectively. The antibody for VP western analysis was a polyclonal rabbit antibody raised against MVM virus. RNase protection assays were performed with homologous probes spanning MVM nucleotides (nt) 1858 to 2377, as previously described (57).
RESULTS:

**APOBEC3A inhibition of MVMp is not at the single strand DNA production step.**

It has been reported that APOBEC3A inhibits replication of AAV2, a member of the parvovirus family with ssDNA genome (9). Since MVM also belongs to parvovirus family and its gene products share a lot of similarity with AAV, we decided to examine the effect of APOBEC3A on MVM replication by cotransfecting APOBEC3A with MVMp infectious clone. In addition to the wild type APOBEC3A, an APOBEC3A mutant, C106S, was also tested. Cysteine 106 resides in a sequence motif $\text{HXEX}_{28}\text{PCX}_{4}\text{C}$ that coordinates a $\text{Zn}^{2+}$ and carries out the cytidine deaminase activity (29). We also included the MVM mutant pMVM(VP1/23451) which does not produce the VP proteins, resulting in the absence of ssDNA production. By analyzing this mutant, we expect to pinpoint the specific step that APOBEC3A exerts its inhibitory effect.

Southern blot analysis shows that APOBEC3A did inhibit MVMp replication (Fig V-2), while other APOBEC3 (hA3B, hA3C, hA3F and hA3G) does not (data not shown). Although the mutant pMVM(VP1/23451) does not produce ssDNA, it is still inhibited by APOBEC3A. Therefore, the effect is not directly through inhibiting the production of ssDNA. In addition, the C106S mutant did not inhibit replication, which indicates the cytidine deaminase domain might be involved in this effect.
Figure V-2. APOBEC3A inhibition of MVM replication is not at the ssDNA production step. MVMp infectious clone and pMVM(VP1/23451) mutant are transfected with empty vector (-), APOBEC3A (3A) and the APOBEC3A mutant (C106S). Marker (M) and wild type virus infection (V) control is also included. Different replication intermediate is labeled as well (D for dimer, M for monomer, ss for single strand DNA).
**APOBEC3A does not affect MVMp RNA or protein production.**

To further investigate the effect of APOBEC3A on MVM life cycle, we decided to examine viral RNA and protein production in the presence of APOBEC3A. As shown in the RNase protection assay in Fig V-3 (A), for both the wild type and mutant MVM, the different species of RNA were maintained at similar levels of production in the absence and the presence of APOBEC3A. C106S also did not affect the RNA production. The protein level examined by western blot reflected the RNA level as both the non-structural NS1/2 and structural VP proteins are expressed at similar levels in the absence and presence of APOBEC3A. These results suggested that the inhibitory effect of APOBEC3A was not upon viral RNA or protein production, but rather it directly targets the DNA replication machinery.
Figure V-3. APOBEC3A does not affect MVM RNA or protein production. (A)

MVM RNA production detected by RNase protection assay following cotransfection with empty vector (-), APOBEC3A (3A) and APOBEC3A mutant (C106S). Different species of RNA are labeled on the right side of the gel. (B) MVM protein production under the same condition as in panel (A) detected by Western blot.
DISCUSSION

From these preliminary data, we observed the inhibitory effect of APOBEC3A on MVM replication. Although C106S mutant loses the inhibiting ability, it remains to be confirmed the direct role of cytidine deaminase activity (CDA) in inhibition, as previous report suggested that APOBEC3A retains its inhibitory effect on intercisternal A particle (IAP) replication even when its CDA is mutated (4). The direct evidence of the involvement of deaminase activity would be sequencing the viral progeny and look for C to T hypermutation. The inhibitory effect is not due to the inhibition of ssDNA production, nor the viral RNA or protein components. Most likely, APOBEC3A may target cellular genes that are required for viral replication, which is a more direct and efficient way if it is responsible for inhibiting ssDNA virus replication in general. On the other hand, it may target cell cycle machinery, as MVM is known to promote cellular S phase status to initiate its replication. It would be interesting to see if APOBEC3A delays or even prevents cells from entering S phase by flow cytometry. Reagents that promote S phase can also be tested to see if they counteract the effect of APOBEC3A. It would also be interesting to examine if APOBEC3A interact with viral capsid proteins and is packaged into virion to deliver its effect as is the case for APOBEC3G (18), although it is less likely considering the small size of MVM.
ii. Leader Sequence Control of Capsid Protein Expression in Canine Parvovirus

INTRODUCTION

Canine parvovirus (CPV) is another member of the autonomous parvovirus family that causes lethargy, vomiting and diarrhea in dogs. It is the general consensus that CPV evolved from feline parvovirus and the host range specificity is determined by residue 93 and 323 of the capsid protein VP2 (8).

The CPV genome, similar to other paroviruses, is organized into two major open reading frames encoding non-structural and structural proteins, driven by P4 and P38 promoters, respectively (Fig V-2). The capsid RNA transcribed from P38 undergoes alternative splicing through a small intron composed of two donors and one acceptor, thus producing VP1 and VP2 proteins.

Our initial attempt to express CPV capsid protein for vaccine development suggests the requirement of the leader sequence upstream of the initiation codon for efficient expression of VP2. This part of the research tries to determine the essential elements within this leader sequence that control VP2 expression and tests at which step they exert the effect.
MATERIALS AND METHODS

**Cells and transfection:**

293 cells were propagated as previously described (40). Transfections, using Lipofectamine and the Plus reagent (Invitrogen, Carlsbad, CA), were performed as previously described (52).

**Plasmid constructs:**

A diagram of the constructs tested is shown in Fig V-4. CPV 2020 is based on pcDNA3 with the CPV capsid gene (nt2020-4558) cloned downstream of the CMV promoter using KpnI and XhoI restriction sites. An HA tag is added at the C-terminus of the gene to allow detection of the protein by anti-HA antibody. CPV 2767 is similar to CPV 2020 except the VP2 gene starts from nt 2767. CPV cDNA is based on CPV 2020 except the intron region (nt2280-2387) is deleted. CPV 2767CM is similar to CPV 2767 except a KpnI-SacII-BamHI polylinker is inserted downstream of the promoter to allow insertion of the fragments from CPV or pBR322. The remaining constructs are all based on this plasmid.

2767 LUS has the large exon (nt2388-2766) inserted into the polylinker upstream of the VP2 gene. 2767 J251 LUS has an extra 251nt long fragment from pBR322 (nt2665-2916) inserted upstream of the large exon to make up the space originally occupied by the small exon. 2767 SUS has the small exon (nt2020-2279) inserted upstream of VP2. 2767 SUS J368 has an extra 368nt long fragment from pBR322 (nt2548-2916) inserted downstream of the small exon to make up the space originally occupied by the large exon. 2767 J251 has only 251nt stuffer sequence from pBR322 upstream of the VP2 gene. 2767 J368
has 368nt stuffer sequence form pBR322. J251 J368 has both stuffer sequences to separate the promoter from the gene exactly the same distance as the two exons do.

**RNase protection assay and western blot:**

36-41 hours post-transfection, cells were washed and collected. 4/5 of the cells were processed for RNase protection assay as previously described (42, 57) and 1/5 of the cells were processed for western blot (57). RNase protections were performed as previously described (42, 57), using homologous anti-sense probes, termed VP2, which spanned nts 2881-3040 of the CPV VP2 region. An monoclonal anti-HA antibody (Sigma, MO) is used for western blot.
Figure V-4. Diagrams of the constructs used in CPV upstream sequence study. All constructs are based on pcDNA3 backbone. Arrow represents the CMV promoter. Exons are shown in rectangular and intron is shown in thin lines. The upstream small exon is represented by the black box while the downstream large exon is in grey box. The small stuffer sequence from pBR322 is in vertical striped box and the large stuffer sequence is in horizontal striped box.
RESULTS

**CPV VP2 upstream exon sequence controls RNA and protein production.**

Previous observations in our lab indicated that VP2 upstream sequence (nt 2020 to nt 2766) was required for efficient expression of the protein in the baculovirus system, even when the gene was driven by a strong CMV promoter (unpublished data). This sequence is composed of two exons and an intron. To narrow down the region that is responsible for this control, series of constructs were made to delete or replace each element. Following transfection into 293 cells, the production of RNA was examined by RNase protection assay and the corresponding protein level of each sample was tested with western blot. Similar to what was observed in the baculovirus system, the full length CPV 2020 construct produced sufficient amount of VP2 RNA while CPV 2767 and CPV 2767CM which lack the upstream sequence did not make VP2 RNA. CPV 2767 cDNA which lacks the intron still produced abundant VP2 RNA (Fig. V-5(A), compare CPV cDNA to CPV 2020), which suggested the intron portion is dispensable. Adding each exon back increased the RNA production, with the smaller upstream exon having a greater effect compared to the larger downstream exon (compare CPV 2767 SUS and CPV 2767 LUS to CPV 2767). By just adding irrelevant sequence upstream of the gene to make up the distance occupied by the exons, the RNA level was only slightly increased (compare CPV 2767 J251, CPV 2767 J268 and CPV 2767 J251 J368 to CPV 2767), possibly due to the separation of the promoter from the coding sequence. This indicated
that the exons harbor some specific elements augmenting RNA production. However, distance does have an additive effect in addition to these elements as inserting extra stuffer sequences between the promoter and the gene further increased RNA level (compare CPV 2767 J251 LUS to CPV 2767 LUS, CPV 2767 SUS J368 to CPV 2767 SUS). The level of protein production of each construct mirrors the RNA level (Fig V-5(B)), which suggested that the difference we observed took place before the translation step.
Figure V-5. CPV capsid RNA and protein production regulation by different upstream sequences. (A) RNase protection assay of RNA collected following transfection of the constructed illustrated in Fig V-4. GFP construct was cotransfected as transfection control. The quantification result was shown on the right panel using Fuji Multigauge software. (B) Western blot for the same constructs transfected as in panel (A) using anti-HA monoclonal antibody (Sigma-Aldrich, St. Louis).
DISCUSSION

Efficient expression of the CVP VP2 gene is not only required for viral capsid production, but also holds promises for vaccine generation. Our preliminary data suggests that the upstream sequence of VP2 is required for its efficient expression, acting at the RNA level. It could either increase the transcription of the gene, or slow down the degradation of the RNA. Although the mechanism remains to be discovered, it would be interesting to see first if the same phenomenon happens with the native CPV promoter in the natural CPV host cells. We are currently testing this using CPV P38 driven VP2 constructs in CRFK cells, which is a feline cell line supporting CPV replication. However, due to the low basal level of P38 promoter, the non-structural protein NS1 is required to activate the promoter to yield sufficient amount of signal to be observed.

Further investigation can be carried out to identify the specific elements in the sequence that augment transcription or prolong RNA half-life.
VI. SUMMARY

Eukaryotic gene regulation is a complicated process that involves vast
collection of cellular factors. Viruses hijack such machinery for their own gene
expression, usually involving modulation by viral factors. For viruses with
compact genomes such as paroviruses, precise regulation of gene expression
in accordance with viral life cycle is especially essential. In this study, we
systematically compared the transcription of the capsid promoters from adeno-
associated virus serotype 2, 5 and bovine adeno-associated virus, and
discovered distinct strategies employed by these viruses despite their sequence
homology.

For AAV2 and AAV5, the differences in capsid promoter regulation are at
both the trans activator and cis sequence level. AAV2 P40 transcription level is
low and relies on its non-structural large Rep protein for activation in the
presence of helper viral functions. This activation is proposed to be mediated by
a loop structure formed via the interaction between AAV2 Rep78 and
transcription factor SP1 and their perspective binding sites in ITR/P5 and P41
region (46) (Fig. VI-1). This structure helps to recruit transcription machineries
from ITR/P5 to the capsid promoter region, hence increasing P40 transcription.
AAV5 P41 promoter, on the other hand, does not require Rep activation as its
basal level is constitutively high in the E1A/B expressing 293 cells. The higher
basal activity of AAV5 P41 promoter is at least partially contributable to the AP1
and CRE transcription factor binding sites upstream. The domain governing the
activation difference of the Rep protein resides in the amino terminus of the
Figure VI-1. Summary of capsid promoter regulation of various adeno-associated viruses. Arrows represent AAV2, AAV5 and BAAV capsid promoters. Filled ovals represent Rep binding element (residing in ITRs at the end of each genome and an extra one in P5 region for AAV2). Adenovirus genes are shown in a cluster of empty ovals. The upstream AP1 and CRE transcription factor binding sites are in rectangular and labeled with “+”. The top panel is derived from previous study by Pereira et al (46). The rest is based on this study. See details in text.
protein. Although closely related to AAV5, BAAV demonstrates a transcription profile similar to AAV2. Despite the presence of AP1 and CRE sites upstream, the basal level of P41 promoter is low due to the inhibitory effect of surrounding sequences. However, it overcomes this effect with the activation of the promoter by its Rep protein, possibly through a mechanism similar to AAV2.

These studies mainly focus on the viral transcription control on the molecular biology level, but it also revealed possible evolutionary relationships between these viruses. AAV5, as the most divergent serotype among human AAVs, exhibits a distinct transcription map. We found that it is more closely related to AAVs from animal origins, not only on the sequence homology level, but also in terms of transcription modulation, possibly indicating a stage of host transition. This knowledge will benefit future research using viral genes as models to study eukaryotic transcription, and it also provided a tool to trace host migration of closely related viruses.
Reference


VITA

Chaoyang Ye was born to Aimin Li and Xiayu Ye on February 3, 1979 in Zhenjiang, Jiangsu, China. He received his Bachelor of Science degree in June, 1997 from Nanjing University, China and came to the United States in August, 1997 for graduate study. Chaoyang completed his Doctorate of Philosophy in Molecular Microbiology and Immunology from the University of Missouri-Columbia in May, 2007.