[Frontiers in Bioscience 13, 3101-3115, January 1, 2008]

Processing of adeno-associated virus RNA

Jianming Qiu[2], David Pintel[1]

[1]Department of Molecular Microbiology and Immunology, University of Missouri-Columbia, School of Medicine, Life Sciences Center, Columbia, MO 65212, [2]Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Kansas City, KS 66160

TABLE OF CONTENTS

1. Abstract
2. Introduction and General Considerations
3. AAV2-like adeno-associated viruses
- *3.1. General transcription organization of AAV2*
- *3.2. AAV2 RNA splicing*
  - *3.2.1. Helper virus function in AAV2 RNA splicing*
  - *3.2.2. The role of the large Rep proteins in AAV2 RNA splicing*
  - *3.2.3. The size of the AAV 5' exon is inversely correlated with splicing enhancement by Rep*
  - *3.2.4. Splicing enhancement by Rep, in the presence of adenovirus, is maximal in conjunction with the AAV2 transcription unit*
- *3.3. RNA splicing of other AAV2-like serotypes of AAV*
- *3.4. Export of AAV2 RNA*

4. AAV5-like adeno-associated viruses
- *4.1. Transcription profile of AAV5*
- *4.2. Splicing of AAV5 RNA*
- *4.3. Polyadenylation of AAV5 RNA*
  - *4.3.1. Internal polyadenylation*
  - *4.3.2. cis-Determinants of AAV5 RNA internal polyadenylation*
  - *4.3.3. Regulation of AAV5 internal polyadenylation*
  - *4.3.4. 5'-terminal exon definition model*
- *4.4. Bovine adeno-associated virus (B-AAV) and Goat adeno-associated virus (AAV-Go.1)*

5. Avian-adeno-associated virus
6. Concluding remarks
7. Acknowledgements
8. References

1. ABSTRACT

Adeno-associated viruses (AAVs), members of the Dependovirus genus of the subfamily *Parvovirinae*, family *Parvoviridae*, have single-stranded DNA genomes that replicate *via* a double-stranded DNA intermediate which serves as the transcription template. They exhibit a highly compact, overlapping genetic organization which is surprisingly variable amongst various members within the group. Additionally, recent work has shown them to be excellent models for understanding potential interactions between alternative splicing and alternative polyadenylation, the export of unspliced RNAs, and transactivating proteins that influence RNA processing co-transcriptionally via interaction with the transcription template.

2. INTRODUCTION AND GENERAL CONSIDERATIONS

The parvoviruses are a group of small non-enveloped T=1 icosahedral viruses approximately 18-26 nm in diameter (1,2). They are unique among all known viruses in that they contain single-stranded linear DNA genomes (1-3). Members of the family *Parvoviridae* that infect vertebrate hosts are grouped within the subfamily *Parvovirinae,* which is divided further into five genera: *Parvovirus*, *Dependovirus*, *Erythrovirus*, *Amdovirus*, and *Bocavirus* (4). The adeno-associated viruses (AAVs) are members of the *Dependovirus* genus, and so named because its members replicate efficiently only in the presence of helper DNA viruses such as adenoviruses (Ad)

or herpesviruses (HSV). Parvovirus genomes vary between 4-6 kb of single-stranded DNA and their termini contain palindromic sequences (3,5). For members of the *Dependoviruses,* and the Erythrovirus human parvovirus B19, the two terminal hairpin sequences are identical, while for others they are different, however, in all cases, they can form hairpin structures that comprise the viral origins of replication (3,5). Subtle differences in their replication mechanisms result in the packaging of either predominately negative sense DNA strands for members of the *Parvoviruses and Amdovirus genera*, or both positive and negative strands (in different capsids), at either equal levels for members of the *Dependovirus* and *Erythrovirus* genera, or intermediate levels for members of the *Bocaviruses* genera (1,3,5).

AAVs have been isolated from multiple species including humans and other primates (6-10), birds (11), bovine (12), sheep (13), snake (14) and rodent (15). As described more fully below, they all use three promoters to generate the spectrum of transcripts needed to encode the viral proteins required for infection, and they all feature a single intron with a single donor and two acceptors in the center of their genome. All AAVs encode at least two nonstructural Rep proteins, and three capsid proteins from open reading frames (ORFs) in the left- and right-hand halves of genome, respectively. The early characterization of the transcription profile of AAV2 has appeared in numerous reviews (16-18). In this manuscript we review more recent results that include analyses of various AAV serotypes.

Survey of the transcriptional profiles of primate AAV serotypes 1-6, and other animal AAVs has shown that they fall into three main groups. The first group, for which AAV2 can be taken as the prototype, includes AAV1, AAV2, AAV3, AAV4, and AAV6. These viruses all possess similar large introns (approximately 320 nts) which feature non-consensus donor and acceptor sites (Figures 1&2). Their RNAs are all are polyadenylated exclusively at a site at the right hand end of the genome (Figure 1). Interestingly, as discussed further below, the large Rep proteins of these viruses also participate in viral RNA processing, but to different degrees amongst the various serotypes. Although they are all dependent upon larger DNA viruses for replication, the different serotypes in this group also shows a variable dependence on their helper viruses for expression. AAV2, AAV3 and AAV4 are very dependent on both helper virus and Rep for proper RNA processing and high levels of transcription initiation, while AAV1 and AAV6 generate relatively lower levels of spliced RNA products and are significantly less responsive to Ad5 helper functions for enhancement of RNA processing (J. Qiu, D. Farris and D. Pintel, unpublished).

The second main group of AAVs comprises AAV5 and the animal AAVs, Bovine AAV (B-AAV) and Goat AAV (AAV-Go.1), and possibly the rat and mouse AAVs (15). They utilize a significantly different expression strategy: RNAs generated from the two viral upstream promoters utilize an internal polyadenylation site within the central viral intron, while RNAs generated by the promoter-proximal capsid gene promoter read through this site and are polyadenylated at the site at the 3' end of the genome and subsequently spliced (Figure 3) (19-21). Plasmid clones of these viruses express constitutively higher levels of spliced capsid-encoding RNAs and so are less dependent on helper virus gene products or Rep for their expression than the AAV2-like subgroup (22-26).

The avian AAV (A-AAV) falls into the third group. Its expression pattern seems to be a hybrid of AAV2 and the AAV5 strategies. A-AAV utilizes an internal polyadenylation site within the central intron, however, at a relative low efficiency (20). Approximately 50% of A-AAV RNAs generated by its upstream promoters are polyadenylated at the internal site, and those that read through accumulate as both spliced and unspliced species (20).

Portions of the genomes of AAV7, AAV8, AAV11 and AAV12 have been isolated from monkey tissue (10,27), and mouse and rat AAVs (15) have also been isolated from rodent tissues, however, these have not yet been characterized as infectious virus. The introns of AAV7 and AAV8 are similar to the AAVs of the AAV2 subtype and so it is likely that their RNA profiles will be similar to that group. The murine-AAV and rat-AAV introns are more similar to the AAV5 and animal AAV subgroup, and so their RNA profiles may resemble that group.

AAV RNA processing is an excellent model to study alternative RNA processing of complex overlapping transcription units in a small but complete biological system. These investigations have provided insight into the basic relationship between splicing and polyadenylation, and into the co-transcriptional nature of these processes. Thus far, the only parvovirus protein that has been shown to participate in its efficient RNA processing is the AAV2 Rep78/68 protein. Whether its role is direct, or facilitates the function of Ad gene products in this regard is not known, and will be discussed further below.

Although the characterization of parvovirus alternative splicing and polyadenylation has been extensive, very little is known about the export of AAV RNAs. Interestingly, two critical Rep proteins of AAV2 (Rep78 and Rep52), and the Rep proteins of AAV5 are generated from completely unspliced RNAs. How these RNAs, which are not processed by the splicesome, are exported to the cytoplasm at high efficiency is currently unclear.

3. AAV2-Like ADENO-ASSOCIATED VIRUES

3.1. General transcription organization of AAV2

The transcriptional organization of AAV2 has been previously reviewed (17,18). The AAV2 genome has three promoters identified by their map positions of P5, P19, and P40 (28,29) (see Figure 1). All of the AAV2 transcripts contain a single intron, which uses a 5' splice donor at nt 1906, and one of either two splice acceptors at nt 2201 (A1) or nt 2228 (A2) (see (17)). Unspliced P5-

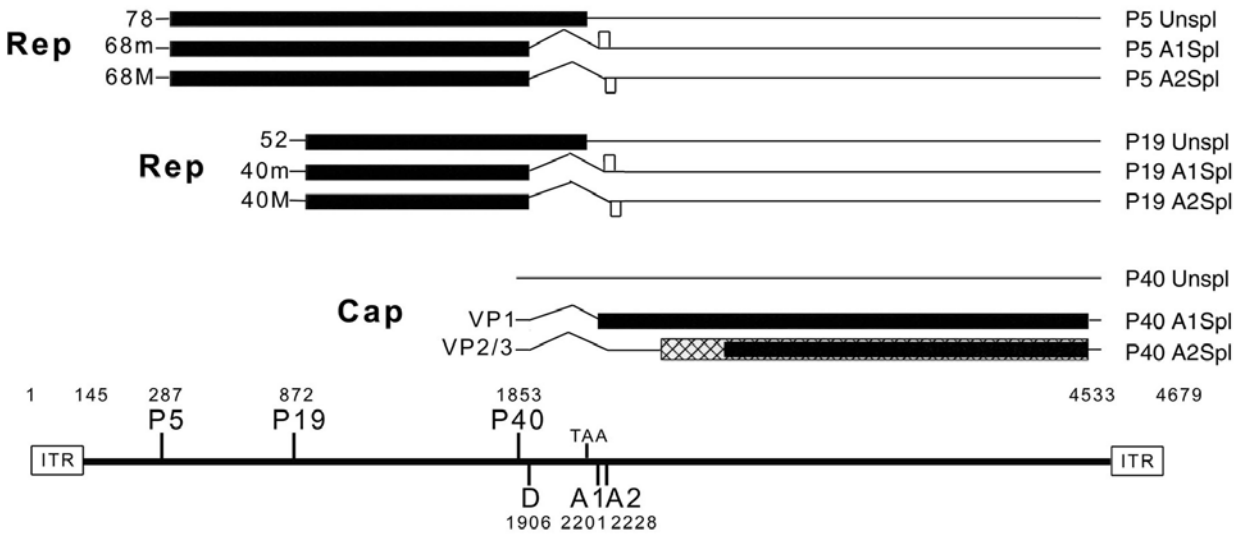

Figure 1. Transcription map of the AAV2. The AAV2 genome of 4,679-nt is depicted at the bottom of the figure. The inverted terminal repeats (ITR) are shown. All the transcription landmarks, the P5, P19 and the P40 promoter, the donor (D) and two acceptor sites (A1 and A2) of the central intron are shown with designations of their respective nucleotide numbers, which are referred to the prototype AAV2 sequence (Genebank accession number: AF043303). All the transcripts from the three AAV promoters are shown with their designations at the right, and the proteins that they encode are shown with designations at the left. m represents minor transcripts spliced from A1 acceptor, while M represents the major transcripts spliced from the A2 acceptor.

and P19-generated RNAs encode Rep 78, and Rep 52, respectively, while spliced P5- and P19-generated RNAs encode Rep 68 and Rep 40. The large Rep proteins are required at various stages of both genome replication and expression (reviewed in (16)), while the small Rep proteins appear to have critical roles in viral packaging (30,31). Alternative splicing to either A1 or A2 generates two isoforms each for Rep 68 (Rep 68m and Rep 68M) and Rep 40 (Rep 40m and Rep 40M) (Figure 1), however, the importance of this variation is not clear. Spliced P40-generated RNAs utilizing the A1 acceptor encode the VP1 capsid protein, while spliced RNAs utilizing the A2 acceptor encode VP2 and VP3 (32). The relative usage of A1 and A2, which thus determines the relative level of VP1 vs VP2/ VP3, is precisely regulated at a ratio of 1:10 at late times following infection, presumably by the relative efficiency of the two 3' splice sites. Unspliced P40 mRNAs exit to the cytoplasm, however, they express very little capsid proteins, even though they contain the entire VP1 open reading frame (33). This may be because there are a number of short ORFs within 5' UTR of the unspliced P40 transcripts that are available (17). Recombinant AAV vectors have been shown to express transcripts generated from the AAV2 ITR (34-36), and a similar set of ITR-initiated transcripts have been detected at low levels during AAV2 infection (R. Nayak, G. Tullis, and D. Pintel, unpublished). The biological function of these ITR-initiated transcripts remains undetermined. As mentioned, AAV2 gene expression is tightly controlled by both the AAV-encoded Rep proteins (26) and co-infecting helper virus (see (17)).

Similar to the autonomous parvovirus minute virus of mice (MVM), there is a temporal order to the appearance of AAV2 RNAs during infection (24,32,37). Unspliced P5-generated transcripts are detectable prior to the significant accumulation of other AAV2 RNAs. Ultimately, P19-generated RNAs accumulate to levels greater than those generated from P5, and at late times post-infection, P40-generated transcripts come to predominate in the total RNA pool. In addition, the percentage of each class of AAV2 RNA that is spliced increases during infection (24). The degree of this increase is different for the P5 and P19 RNA compared to those generated by P40. At late times post infection, 90% of P40 RNAs, but only 50% of RNA generated by P19, and less than 10% of RNAs generated by P5, are spliced (24). Thus, surprisingly, the same AAV2 intron is excised to different final levels from these different RNA species. It has also been observed that the ratio of splicing to A2 relative to A1 increases as AAV2 infection progresses, however, as noted above, the functional relevance of this is not clear (24). All AAV2 RNAs have been shown to be quite stable during infection (24).

3.2. AAV2 RNA splicing

Pre-mRNA splicing is a critical determinant for productive AAV infection. The ratio of the viral capsid proteins, as well as the ratio of Rep 78 and Rep 52, *versus* Rep 68 and Rep 40 proteins depends upon the level of splicing of the various AAV RNAs (17,18). Splicing of AAV2 pre-mRNAs requires the participation of both the helper virus (e.g. adenovirus (22-25) or herpes virus (24,38,39)), and the large Rep proteins (26). As mentioned above, the majority of P40-generated RNAs are spliced in the presence of adenovirus, however the steady-state ratios of spliced to unspliced RNAs generated by the P5 and P19 promoters are much less.

3.2.1. Helper virus function in AAV2 RNA splicing

Co-infection of cells with AAV2 and either adenovirus or herpes virus stimulates the splicing of AAV2 RNA, but this is manifest only in the presence of Rep (26).

In the absence of either helper virus or Rep, splicing of both P40 and P19-generated transcripts produced following plasmid transfection is very poor (24,40).

Adenovirus clearly has multiple roles in supporting AAV2 infection (22,40,41). Five adenovirus gene products (E1A, E1B, E2a, E4orf6 and VA RNA) have been shown to be required for productive AAV2 replication and recombinant virus production (42-47). 293 cells, which express the adenovirus E1A and E1B genes, provide the necessary helper environment for AAV2 replication when transfected with a plasmid that expresses the adenovirus gene products E2a, E4 and VA (e.g. pXX6) (40,48). The splicing of AAV2 RNAs in 293 cells following transfection of either an infectious viral clone, or a non-replicating, ITR-minus AAV2 RepCap plasmid, is stimulated by by pXX6 as effectively as adenovirus co-infection (J. Qiu and D. Pintel, unpublished); however, expression of either E2a, E4, or VA genes individually is not adequate (J. Qiu and D. Pintel, unpublished). This suggests that Ad gene products stimulate splicing of AAV RNA independent of their effect on replication, however, how these adenovirus gene products function, in conjunction with Rep, to enhance AAV RNA splicing is not clear. Additionally, the herpes virus UL5, UL8, UL52 and UL29 gene products can fully support AAV2 replication (49). That successful recombinant AAV production can be achieved using those herpes gene products suggests that they can, along with Rep, also efficiently enhance splicing of the P40-generated RNA. However, a direct role for these gene products in enhancement of AAV2 RNA splicing has not yet been shown. Comparison of the role of adenovirus and herpes virus gene products in AAV2 RNA splicing should reveal interesting insights into this aspect of their helper function.

3.2.2. The role of the large Rep proteins in AAV2 RNA splicing

In the presence of helper virus gene products, the AAV2 large Rep proteins (Rep 78/68) can act to increase the ratio of spliced to unspliced AAV RNA when they are targeted to the transcription template via a Rep binding element (RBE) (26). An extended RBE, rather than merely the core RBE, provided either by the P5 promoter or the viral ITR is both necessary and sufficient to target Rep in a manner which sustains full enhancement of the splicing of P40-generated RNA; if the transcription template lacks an extended RBE, neither Ad5 nor Rep have significant detectable effects on AAV2 RNA splicing (26). Why an extended RBE is required is not known. It has so far not been possible to separate Rep enhancement of splicing from Rep activation of transcription initiation (26). In addition, the required extended RBE is both location and orientation independent, similar to the classically defined transcription enhancer-factor binding sites, which suggested that Rep might be acting co-transcriptionally to affect RNA processing (26). Although the Rep effect on AAV2 RNA processing appears to be linked to its transcriptional transactivation activity, transactivation of transcription initiation is not sufficient for enhancement of RNA processing (26). Activation of the P40 transcription unit by other activators targeted to transcription template (such as Gal4-VP16, and MVM NS1) does not enhance AAV2 RNA splicing, although transcription of the P40 promoter can be strongly transactivated by these proteins. These results suggest a model in which Rep may act as a trans-regulator of RNA processing by specifically modulating such functions coupled to RNA polymerase II (RNA pol II) transcription (50).

3.2.3. The size of the AAV 5' exon is inversely correlated with splicing enhancement by Rep

As mentioned above, the ratios of unspliced to spliced P5- and P19-generated RNAs is significantly less than for RNAs generated by P40. Similarly, Rep enhancement of the relative levels of spliced RNA decreases as the distance between the promoter and the intron of the affected transcription unit increases (26). This is consistent with a co-transcriptional model of enhancement of RNA processing by Rep, and may explain why the same AAV2 intron is excised to a much greater extent from the proximal P40-generated transcripts than from upstream P5 and P19 transcripts (26).

It is not yet clear how Rep and adenovirus effect AAV RNA processing, indeed, it is not yet clear whether the combined effect of adenovirus and Rep is directly on splicing or other RNA processing mechanisms that ultimately determine the steady-state levels of spliced AAV RNA. Together these factors may alter the composition of the elongating RNA polymerase II complex by altering levels or types of RNA processing factors associated with the RNA polymerase II carboxyl terminal domain (CTD). Alternatively, a direct interaction of Rep with cellular splicing factors could be possible. The recent demonstration that Rep co-localizes in replication centers with helper adenovirus or HSV single-strand DNA binding proteins (ssDBP and ICP8, respectively) following co-transfection (51), suggests it is also possible that Rep functions to enhance AAV2 RNA splicing by targeting AAV2 to these replication centers, which normally are highly active for splicing (52-54).

As described above, the ratio of spliced to unspliced P5- and P19-generated RNAs increases over the course of infection. While the reason for this is not clear, conceivably, this could be due to an increase in Rep production, or perhaps, differences in viral localization as infection proceeds.

3.2.4. Splicing enhancement by Rep, in the presence of adenovirus, is maximal in conjunction with the AAV2 transcription unit

Both the relative constitutive levels of spliced RNA and the ability of Rep, when targeted to the transcription template, to enhance the ratio of spliced to unspliced AAV RNA can vary depending upon the promoter used to generate that RNA. For instance, when the AAV2 promoter is replaced by the HIV LTR (which contains a Rep binding site (55,56)), it generates an approximately equal ratio of spliced to unspliced RNA in the absence of Rep, and the addition of Rep via co-transfection enhances the level of spliced RNA only threefold, compared to the six- to sevenfold enhancement

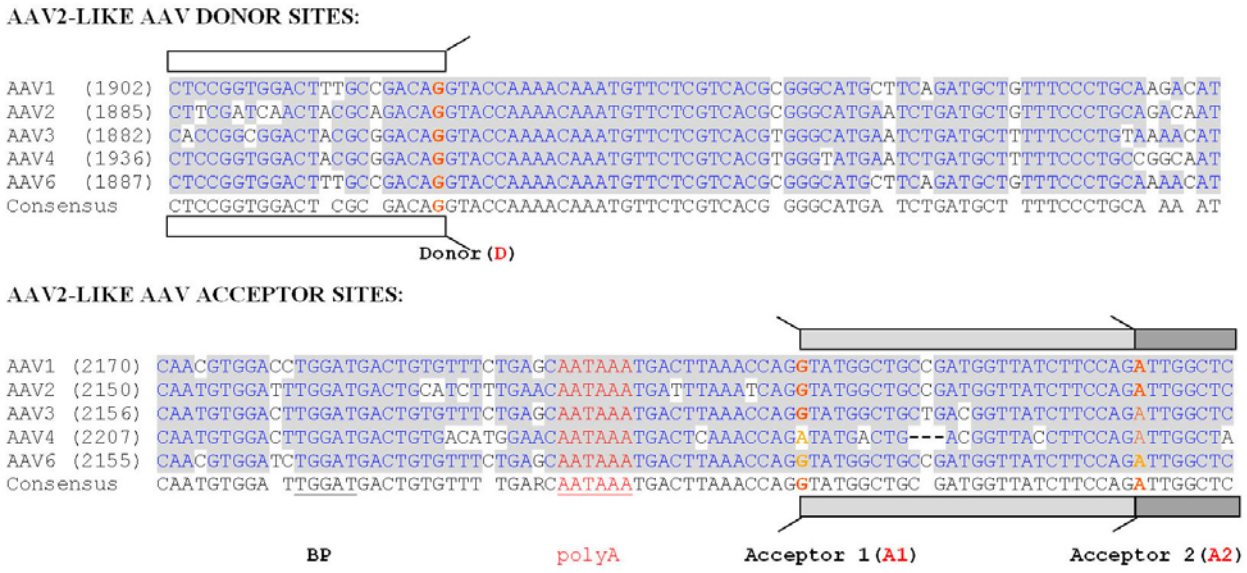

Figure 2. Comparison of the nucleotide sequences of the donor and acceptors sites of the AAV2-like AAVs. The nucleotide sequence of the 5' splice donor region and the 3' splice acceptor are aligned for AAV1, AAV2, AAV3, AAV4 and AAV6. A consensus sequence is presented at the bottom. The intron bordering nucleotides are shown in red. The cryptic polyA signal and the branch point (BP) are underlined.

seen when the capsid gene transcription unit is driven by AAV2 P5 or P40 promoters. When driven by the CMV IE promoter, the ratio of spliced to unspliced RNA from the AAV2 capsid-coding region in the absence of Rep is lower than seen for RNAs generated from either AAV P5, AAV P40 (containing either the P5 or ITR as Rep-binding element), or the HIV LTR. Furthermore, targeting of Rep to the CMV-driven AAV transcription unit only enhances the ratio of spliced to unspliced AAV2 RNA twofold. That the basal ratios of spliced to unspliced AAV RNA differs depending on the promoter used to drive their expression may have been expected from results obtained in other systems (57); however, in addition, the ability of Rep to enhance the levels of the final spliced product is also promoter specific (26).

Both the AAV intron and the extended polyadenylation site are required to retain full responsiveness to splicing enhancement by Rep. The most efficient enhancement by Rep occurrs when an extended sequence including AAV sequences upstream of the canonical polyadenylation site (nt 3941 to 4492) are included in the transcription unit. The core AAV polyadenylation sequence (nt 4326 to 4492), either adjacent to or at various distances downstream of the intron, is unable to sustain Rep responsiveness. These results demonstrate that the effect of the Rep protein in influencing the ratio of spliced to unspliced RNA is not universal and that it is most effective when targeted to an AAV2 transcription unit containing the AAV2 intron paired to its natural extended polyadenylation site and, for maximum effect, driven by the AAV2 P5 or P40 promoters. Why these cis-acting signals respond best to Rep enhancement will likely shed important light on the mechanism its action.

3.3. RNA splicing of other AAV2-like serotypes of AAV

AAVs 1, 2, 3, 4, and 6 share substantial homology within their intron regions, and exhibit identical donor sites and highly similar acceptor sites (see Figure 2). Splicing of AAV3 and AAV4 is basically similar to that of AAV2 with respect to its efficiency and dependence on helper virus and Rep, however, the splicing of AAV1 and AAV6 appears to be less dependent on helper virus (J. Qiu, D. Farris and D. Pintel, unpublished). Most of the P41-generated transcripts generated by AAV1 and AAV6 RepCap plasmids following transfection of 293 cells remain unspliced even in the presence of Ad5. Lower levels of splicing of P40-generated products are also seen during AAV1 and AAV6 viral infection. The levels of splicing seem to be governed by intron sequences. The AAV2 intron, when placed in an AAV1 background, splices at high efficiency, while the AAV1 intron still splices poorly when inserted into AAV2 (J. Qiu, D. Farris and D. Pintel, unpublished). These results suggest that it may be possible to identify *cis*-acting sequences within the AAV2 intron that confer splicing responsiveness to adenovirus and Rep.

3.4. Export of AAV2 RNA

An unusual feature of AAV2 gene expression is that unspliced RNAs from the P5, P19 and P40 promoters accumulate in the cytoplasm at high levels. The P5- and P19-generated unspliced RNAs are translated to generate the non-structural proteins, Rep 78 and Rep 52, respectively. Typically, unspliced cellular and viral RNAs remain in the nucleus (58). Their export usually requires a regulated pathway including a specific *cis*-acting export signal, such as the CTE of the HIV rev RNA (59). The export and translation of unspliced P5 and P19 generated

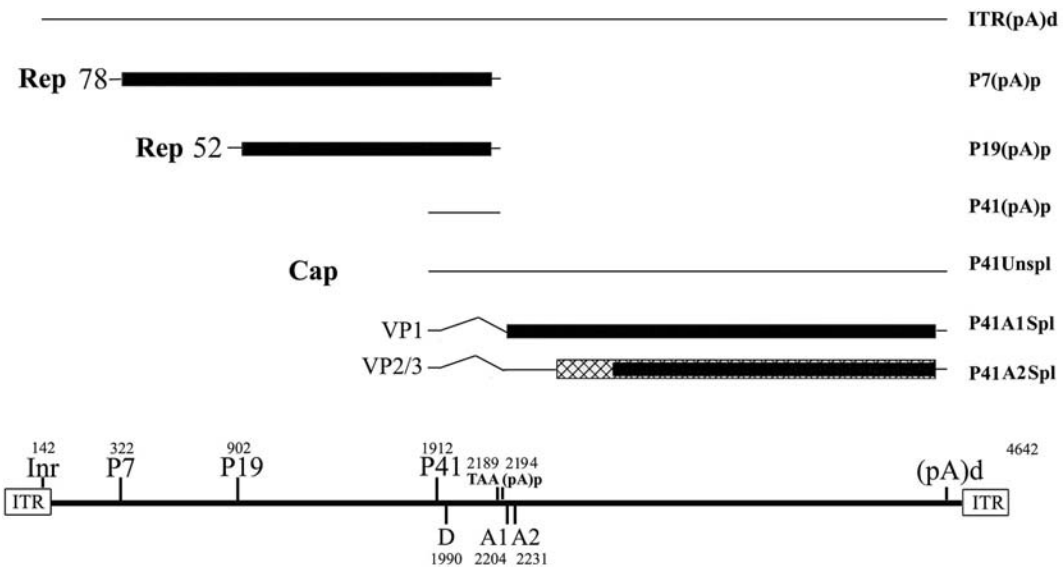

Figure 3. Transcription map of the AAV5. The 4,642-nt AAV5 genome is shown to scale with the major transcription landmarks, including the ITRs, promoters, initiator (Inr) site within the ITR, the initiation sites for the various RNAs, splice donor (D) and acceptors (A1 and A2), and the proximal ((pA)p) and distal ((pA)d) polyadenylation sites. Their nucleotide numbers refer to the AAV5 sequence GenBank accession number: NC_006152. The major transcripts observed are listed on the right and the proteins that they encode are shown on the left. As mentioned in the text only two Rep proteins are encoded. The cap open reading frames (ORFs) that are used are shown in different shading patterns.

RNAs may compensate for the poor splicing of P5- and P19-generated RNAs, impaired due to their distance from the intron (see above).

During adenovirus infection the adenovirus E1B55K-E4orf6 complex both inhibits the export of cellular mRNAs, and is required for efficient cytoplasmic accumulation of adenovirus mRNAs (60-64). The E1B55K-E4orf6 complex has also been suggested to be necessary for the timely cytoplasmic accumulation of AAV RNAs during adenovirus co-infection (44). These experiments were performed using adenovirus mutants that individually did not generate either the E1B 55K or E4orf6 proteins. Because of recent experiments highlighting the critical importance of these proteins in multiple steps during the replication and expression of AAV2, their role in export of AAV2 RNA likely needs to be re-evaluated. Following transfection of an infectious clone of AAV2 into 293 cells in the absence of Ad5, unspliced AAV2 RNAs can be efficiently exported to the cytoplasm (24), suggesting that their export can be achieved in the absence of additional adenovirus gene products.

Whether the export of unspliced AAV2 RNAs requires a particular *cis*-acting signal that directs it through a regulated export pathway, or, alternatively, whether these RNAs are exported via the default cellular mRNA export pathway will be important to determine.

4. AAV5-LIKE ADENO-ASSOCIATED VIRUSES

4.1. Transcription profile of AAV5

The basic transcription profile of AAV5 has similarities to that of AAV2, but there are also striking differences (see Figure 3). The main difference between the AAV5 and AAV2 transcription maps is that RNAs generated by the AAV5 P7 and P19 promoters predominately polyadenylate at a site ((pA)p, at nt 2194) within the intronic region of the genome (19) (nucleotide numbers refer to GenBank accession number: NC_006152). Inspection of the intron region of AAV5 shows the presence of two relatively consensus AAUAAA signals at nt 2177 and 2191, which are immediately upstream of the first intron acceptor A1 at nt 2205. RNA cleavage and polyadenylation occur 11 to 14 nt downstream of the first AAUAAA motif, within the second AAUAAA (19). P41 RNAs utilize the (pA)p site with significantly reduced efficiency (65). Two unspliced P41-driven species are detected, a minor one polyadenylated at (pA)p and the major one polyadenylated at the distal site (pA)d at nt 4434. P41-generated RNAs polyadenylated at (pA)d are also spliced, using the A1 or A2 acceptor at a relative ratio of approximately 1:10, which generate VP1 and VP2/VP3, respectively. The function of the unspliced P41-generated RNAs is unknown. A consequence of the internal polyadenylation of the P7- and P19-generated RNAs is that only two Rep proteins, Rep78 and Rep52, are generated during infection.

Another significant difference between the AAV5 and AAV2 transcription maps is that in 293 cells, AAV5 generates an abundant transcript from its ITR (19). This ITR- generated RNA, which extends to the right-hand end of the genome, uses an initiation site that maps to the AAV5 ITR terminal resolution site (TRS), which in AAV5 is similar to a consensus transcription initiator sequence (5'Y-Y-A-N-T/A-Y-Y-Y3' (where Y is C/T and N is any nucleotide)) (66). Although this 4.3-kb RNA is polyadenylated at the distal (pA)d site, it does not appear to

be spliced (J. Qiu and D. Pintel, unpublished). Different accumulated levels of this ITR- generated transcript were observed under different experimental conditions, but at its most abundant, it is found to be nearly as prevalent as the P7-generated transcript (19), making it significantly more abundant that the ITR-generated product previously identified for AAV2 (36). Whether this RNA encodes a protein product, and what role it may play during viral infection, is currently unknown.

4.2. Splicing of AAV5 RNA

The AAV5 intron, which is 240 nts, is somewhat smaller than the 322 nts AAV2 intron. It has a non-consensus 5' splice site (AAG/GTATGA) at nt 1990, and two 3' acceptor sites, A1 and A2, at nt 2204 and 2231, respectively, (please see Figure 3) (19). The polypyrimidine tract in the acceptor region is pyrimidine-rich, and splicing of the AAV5 intron is constitutively high even in the absence of adenovirus (19). Interestingly, splicing at the A1 site is dependent on the A2 site; mutation of the A2 AG motif also debilitates splicing to A1 for reasons that are not yet clear (J. Qiu, C. Ye, and D. Pintel, unpublished). This is not the case for AAV2, for which splicing to A1 is independent of splicing to A2 (J. Qiu, C. Ye, and D. Pintel, unpublished).

In contrast to AAV2, neither the Rep protein nor additional adenovirus gene products are required to achieve efficient AAV5 promoter activity and pre-mRNA splicing following transfection of a RepCap plasmid clone lacking ITRs into E1A- and E1B-containing 293 cells (19). The basal levels of expression and splicing of AAV5 P41-generated RNA under these conditions are high, and are not further stimulated by Ad5 infection (67). Therefore, the expression of AAV5 is considerably less dependent on helper virus than is that of AAV2. Although the requirement that the AAV5 life cycle has for Ad5 may be more limited, helper virus is still required for AAV5 replication following either infection of AAV5 or transfection of the AAV5 infectious clone (19). The independence of AAV5 splicing is not due merely to the smaller size of its intron, but rather to the nature of the *cis*-sequences within the context of the intron (65), in a way that has yet to be fully defined.

4.3. Polyadenylation of AAV5 RNA

Polyadenylation of AAV5 RNAs at an internal site is a hallmark which differentiates this group from the other AAVs.

4.3.1. Internal polyadenylation

AAV5 RNAs generated from P7 and P19, which encode the large and small Rep proteins, respectively, from the large open reading frame in the 5' half of the genome, utilize (pA)p at high efficiency (Figure 3) (19,65). RNAs generated from the P41 promoter, however, which encode the virus capsid proteins from the large open reading frame in the 3' half of the genome, utilize this proximal site with significantly reduced efficiency – these RNAs primarily read through to the distal site (pA)d (65). Interestingly, RNAs generated from the ITR initiator are mostly unspliced, read through the (pA)p site, and ultimately are polyadenylated at the (pA)d site. The AAUAAA motif at nt 2177, which is immediately upstream of the first intron acceptor A1 at nt 2204 (65), is used for internal polyadenylation. RNA cleavage and polyadenylation occurs 11 to 14 nts downstream of this signal (19). AAV2 also contains a consensus AAUAAA polyadenylation signal in this region (Figure 2); however, its use can only be detected by enhanced PCR techniques (G. Tullis and D. Pintel, unpublished).

Polyadenylation of P41-generated RNAs at (pA)p is significantly increased when they are generated from a replication-competent molecule such as viral infection or an infectious clone, compared to a non-replicating template (19,65). This is not merely due to the presence of the ITRs in *cis*, because ITR-containing templates that cannot replicate due to Rep mutation do not show this increase. Whether the replication process itself is able to alter polyadenylation efficiency of AAV5 P41-generated RNAs is currently unclear. Perhaps replication-associated changes to the AAV5 transcription template affect downstream RNA processing events, as has been suggested in other systems (68).

4.3.2. *cis*-Determinants of AAV5 RNA internal polyadenylation

Sequences within the AAV5 intron are both necessary and sufficient to direct cleavage and polyadenylation of P7 or P19-generated transcripts. Efficient polyadenylation at (pA)p of all AAV5 RNAs requires a U-rich downstream element (DSE, UUUUGUU) that overlaps with the A2 3' splice acceptor site (Figure 4) (65). In addition, efficient polyadenylation of P7-driven transcripts at (pA)p involves an upstream element (USE) that lies prior to the initiation of the P41 transcripts (between nts 1266-1637) (65). Because the USE lies upstream of the start of the P41-generated RNAs, these RNAs are not subject to the additional level of control this element imposes on P7 or P19-driven RNAs. The USE is not required if the DSE is intact; however, it can at least partially compensate for the loss of DSE function (65). Although the relationship between these two elements is not yet fully clear, the USE, which is lacking in P41-generated RNAs, likely functions to ensure that the AAV5 P7- and P19-generated RNAs are polyadenylated at (pA)p at high efficiency.

As mentioned above, the AAV5 DSE-containing A2 3' acceptor region plays a dual role in both polyadenylation and splicing, and there is clearly a complex regulation between these two processes. The DSE overlaps, but is not congruent with, the 3' splice site of the AAV5 intron (Figure 4). The relative strengths of these two signals must allow both for polyadenylation of the upstream promoter-generated RNAs that generate the essential Rep proteins and for efficient splicing to generate the correct ratio of spliced P41-generated RNAs to produce the individual capsid proteins at the appropriate stoichiometry. The relative efficiency of these processes is controlled to a significant degree by the strength of the *cis*-acting signals governing them, and they likely have evolved to program the optimal amounts of viral gene products required for the AAV5 life cycle (65).

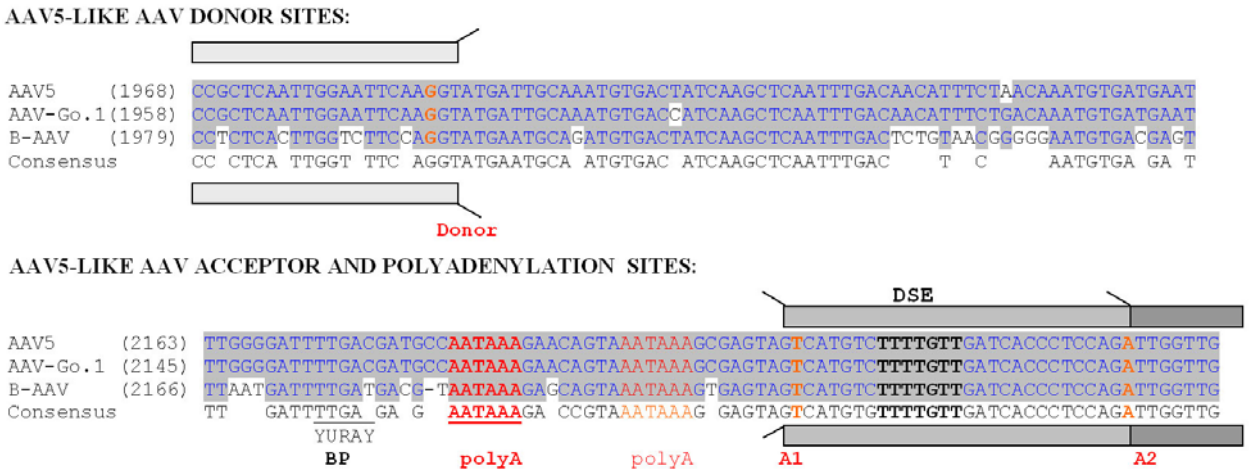

Figure 4. Comparison of the nucleotide sequences of the donor and acceptors sites of the AAV5-like AAVs. The nucleotide sequence of the 5' splice donor region and the 3' splice acceptor are aligned for AAV5, AAV-Go.1 and B-AAV. A consensus sequence is presented at the bottom. The intron bordering nucleotides and the polyA signal are shown in red. The used polyA signal and potential branch point (BP) are underlined.

Interestingly, when the (pA)p region of AAV5 is replaced with the internal AAUAAA signal of AAV2, which is normally cryptic in AAV2, or when the internal region of the AAV2 intron is inserted into AAV5 and linked to the AAV5 DSE-containing A2 3' splice site, polyadenylation at the AAV2 AAUAAA site becomes highly efficient (J. Qiu and D. Pintel, unpublished). This suggests that at least one element that differentiates the use of this site between the two viruses is the presence of the DSE in AAV5. As described more fully below, the DSE region of the avian AAV (A-AAV) (pA)p, which is used inefficiently, is much less consensus than that of the AAV5 (Figure 7). This observation also highlights the importance of this DSE for (pA)p usage.

4.3.3. Regulation of AAV5 RNA internal polyadenylation

Polyadenylation at (pA)p increases as the distance between the RNA initiation site and the intron and (pA)p site is increased (69). The steady-state level of RNAs polyadenylated at (pA)p is independent of the promoter used or of the intervening sequence but is dependent upon competition with splicing, and the intrinsic efficiency of the cleavage/polyadenylation reaction (69). Each of these determinants shows a marked dependence on the distance between the RNA initiation site and the intron and (pA)p.

Apparently, the processing of AAV5 RNA is governed by competition between splicing and alternative polyadenylation for the available pools of AAV5 pre-mRNA (65). Inhibition of internal polyadenylation of AAV5 RNA at (pA)p, required for read-though and eventual splicing of the P41-generated capsid-encoding RNAs, is accomplished by the U1 snRNP binding to the intervening donor in a manner inversely related to the distance between the RNA initiation site and the intron and (pA)p (70). Mutation of the donor in splicing-deficient substrates reduces inhibition of polyladenylation at p(Ap), and introduction of a U1 snRNA gene engineered to rebind the mutant site re-establishes this inhibition (69). The importance of U1 snRNP is also supported by experiments in which introduction of a U-rich stretch, predicted to target TIA-1 and thus increase the affinity of U1 snRNP binding to the intervening donor site, significantly augmented inhibition of (pA)p, while depletion of TIA-1 by siRNA increased (pA)p read-through (70). This suggested that inhibition of polyadenylation of AAV5 RNA at the (pA)p site was governed by the affinity of U1 binding to the intron donor (70). As mentioned above, this competition plays a critical role in the viral life cycle.

4.3.4. 5'-terminal exon definition model

The simplest model to explain how the competition between splicing and alternative polyadenylation in AAV5 is governed is that strong binding of U1 snRNP to the intron donor facilitates both splicing of the intron and inhibition of polyadenylation at (pA)p (portrayed in Figure 5). That polyadenylation at (pA)p is inhibited when the promoter is closer to the intron and (pA)p (favoring read-through and subsequent splicing) and that polyadenylation is favored when the promoter is a greater distance away, are consistent with a model in which U1 snRNP binding is stronger when the promoter is close and is weaker when the distance is large. 5'-terminal exon definition typically requires that factors binding at or near the cap site interact with and stabilize the binding of factors (usually U1 snRNP) to the proximal 5' splice site, and this interaction may be more critical when the proximal 5' donor is nonconsensus (71). The cap structure, via binding to the cap-binding complex (CBC) (72,73) has been suggested to affect both splicing and polyadenylation (72,74). It has been suggested that the nuclear cap-binding complex (CBC) (75,76) facilitates association of U1 snRNP with the cap-proximal 5' splice site, although this interaction has not previously been shown to be strictly dependent on spacing between the cap and the affected

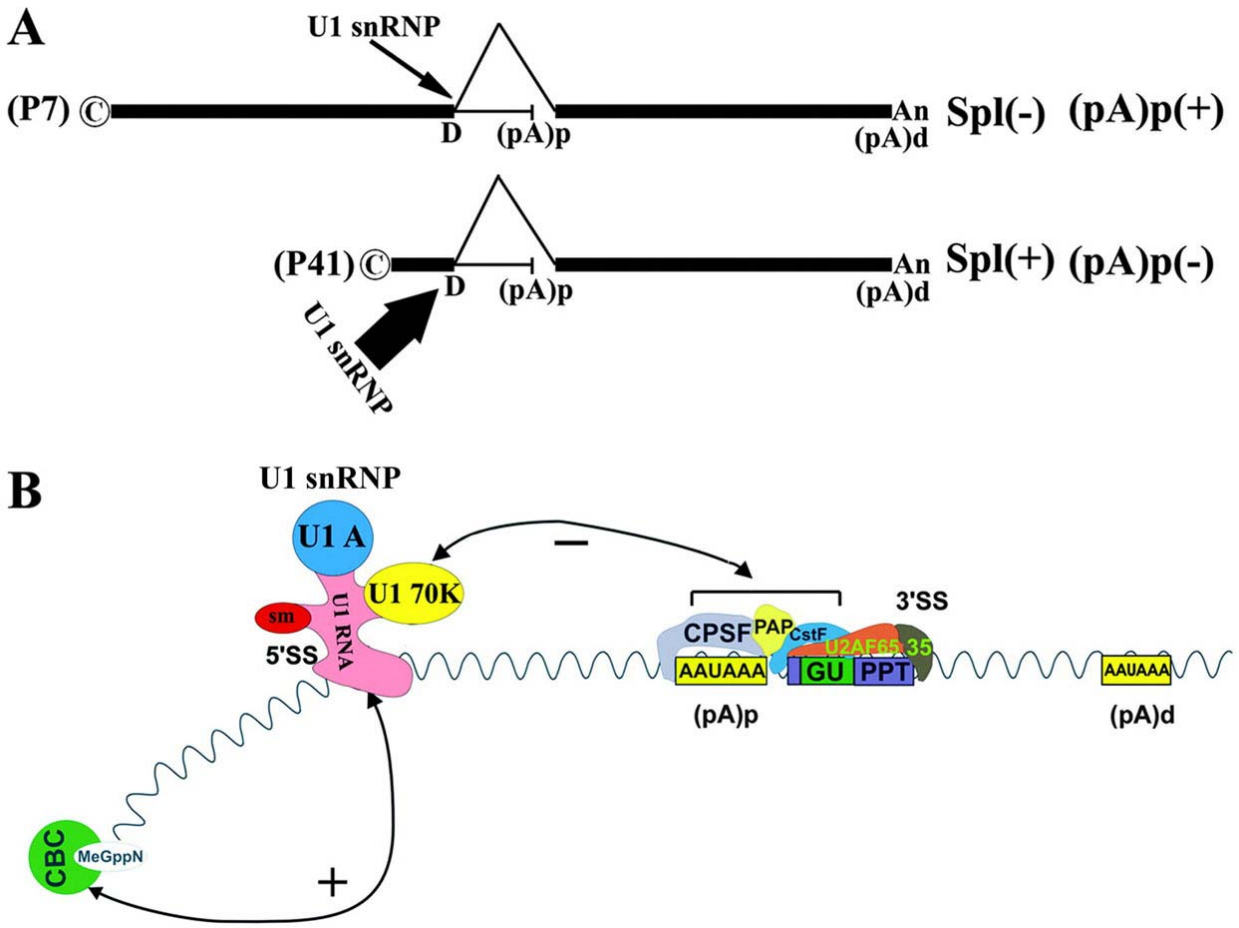

Figure 5. Model for the distance-dependent processing of AAV5 RNA. (A) This diagram shows the processing fate of the AAV5 P7- and P41-generated RNAs. P7-generated RNAs, which have a large 5' exon, are polyadenylated at high efficiency at the internal (pA)p site, which is designated as ((pA)p (+)); few molecules read through and are ultimately spliced, which are designated as (Spl (-)). The opposite is true for P41-generated RNAs. Potential strength of U1 snRNP interaction at the upstream donor, postulated to govern the outcome of the competition between splicing and polyadenylation, is indicated by the width of the arrow. © designates the cap site at the 5'end of the mRNA. (B) This diagram depicts how U1 snRNP binding at the intron donor may exert its negative effect on polyadenylation at (pA)p. The CBC (cap binding complex), the U1snRNP and the major components that interact with the polyA and 3' splice site are diagramed. The U1-70K protein is required for this effect, although the specific target for this inhibition is not yet known. Also shown is the putative binding of splicing factor U2AF(65+35) at the overlapping 3' splice site. Stabilization of U1 snRNP to the intron 5' splice site, which facilitates splicing and simultaneously inhibition of (pA)p, may occur through interaction with the CBC associated with the 5' cap site as shown and may be more effective in this regard when the RNA is initiated closer to the donor site (which promotes splicing and (pA)p inhibition, i.e., with RNA generated from the P41 promoter) than when the RNA is initiated at a distance (such that inhibition of (pA)p is lost, favoring polyadenylation at the internal site, i.e., with RNA generated from the P7 promoter).

'splice site (71). For AAV5, artificially tethering the cap binding complex (CBC) components CBP80 and CBP20 upstream of the intron donor can increase inhibition of polyadenylation at (pA)p (70). The interaction with the CBC seems to strengthen U1 snRNP binding to the downstream AAV5 intron donor in a manner inversely proportional to the size of the 5' exon, thus governing the competition between intron splicing and polyadenylation at (pA)p (70). It is likely that CBC interaction facilitates association of U1 snRNP with the less-than-fully consensus AAV5 pre-mRNA donor site when the cap site is very close (78 nt for the capsid gene promoter P41 (Figure 3)) but less so at a distance (1,668 and 1,088 nt for the large and small Rep gene promoters, respectively). Therefore, the interactions, directly or indirectly, between components of the CBC and U1 snRNP likely mediate the distance related donor inhibition seen for AAV5 due to the strengthening of the association of U1 snRNP to the non-consensus donor site, which would lead to the enhancement of U1's inhibitory activity and govern the competition between intron splicing and polyadenylation at (pA)p (Figure 5). Inhibition by U1 snRNP is mediated by the U1-70K protein binding to loop I on U1 RNA (70). Mutations in the stem-loop I of the U1 RNA that bind the U1-70K protein, but not in the stem-loop II that binds U1-A, are able to abolish inhibition of polyadenylation of AAV5 pre-mRNAs at

(pA)p by U1 snRNP (70). In addition, overexpression of the U1-70K decreases polyadenylation of AAV5 RNA at (pA)p generated following transfection of an AAV5 reporter plasmid in which the distance between the promoter and the intron (with the internal (pA) site) is sufficiently large that polyadenylation of P7-generated transcripts at (pA)p is highly efficient (70).

It is still unknown which part of the cellular polyadenylation complex interacts with the U1-70K protein. U1-70K has been reported to be one of the three U1 specific proteins that binds to a cryptic donor site in front of the distal poly(A) of bovine papillomavirus (BPV), and there it interacts with the poly(A) polymerase, thus controlling processing of BPV late gene RNAs (77,78). However, the situation with AAV5 is likely different. For AAV5, U1 snRNP binds to a functional 5' splice site within the AAV5 intron, thus the U1-70K protein is also involved in normal splicesome assembly. Additionally, unlike previously characterized systems (71,79), inhibition of (pA)p by U1 snRNP binding to the intron donor is decreased as the distance between the donor and (pA)p is increased. This suggests that the interaction may also be influenced by steric effects, and thus, may be more complicated than seen in other systems.

The 5' exon definition model, proposed to regulate the competition between splicing and polyadenylation for AAV5 RNA, is also consistent with the distance-related splicing of AAV2 RNA. In that case, the AAV2 Rep protein enhances splicing of proximal P40-generated transcripts to a much greater degree than it enhances splicing of RNAs generated far upstream from P5. It may be that for AAV2, the Rep protein, in the presence of adenovirus, facilitates the stabilization of U1 snRNP to the AAV2 donor site when the RNA initiation site is close to the intron. A potential mechanism that explains the distance related processing of AAV RNA, using the AAV5 system as a model, is depicted in Figure 5.

4.4. Bovine adeno-associated virus (B-AAV) and Goat adeno-associated virus (AAV-Go.1)

The B-AAV genome is 4,693 nt in length and includes identical ITRs of 173 nt. The Rep protein ORF and the ITRs of B-AAV have high homology to the Rep ORF and ITRs of AAV5 (89 and 96%, respectively) (80). The capsid ORF of B-AAV is most homologous to the capsid ORF of the AAV4 (76%). Goat AAV (AAV-Go.1) was cloned from goat-derived adenovirus preparations (81). Its genome shares 95.2% identity with that of AAV5 and it has identical ITRs (21,82). The Rep ORF shares 99.3% identity with AAV5. However, the C-terminal region of the AAV-Go.1 capsid varies considerably from that of AAV5 (82). A sequence alignment of regions of the donor and acceptor sites of the AAV5, AAV-Go.1 and B-AAV are shown in Figure 4.

In general, the expression profiles of B-AAV and AAV-Go.1 are quite similar to that of AAV5 (19). Like AAV5, B-AAV and AAV-Go.1 utilize an internal polyadenylation site ((pA)p) in the central viral intron efficiently. Similar to AAV5, >99% of B-AAV and AAV-Go.1 RNAs generated from upstream promoters P7 or P19 are polyadenylated at (pA)p and hence not spliced, and therefore, like AAV5, Rep78 and Rep52 are the upstream products primarily expressed (21). Also similar to AAV5, RNAs transcribed from the B-AAV and AAV-Go.1 capsid P41 promoter primarily read through (pA)p and are ultimately spliced (21). The animal derived AAVs thus seem to share with AAV5 the common feature that an internal polyadenylation site is alternatively utilized in their pre-mRNA processing.

There are differences in the level of transcripts generated from the ITRs of these three AAV-5 like viruses. B-AAV produces more ITR-initiated RNA in bovine cells, yet AAV-Go.1 generates little ITR-initiated transcripts in primary goat kidney cells, an observation perhaps not unrelated to differences in cell type. Similar to AAV5, all the transcripts initiated from the ITR of B-AAV read-through the internal (pA)p site and are all polyadenylated distally at (pA)d. It is unknown whether RNAs initiated from the B-AAV ITR is spliced or not in the central intron.

In addition, like AAV2, transactivation of transcription of the capsid-gene promoter of B-AAV requires both adenovirus and targeting of their Rep proteins to the transcription template. In contrast to AAV2, however, splicing of the B-AAV P41- generated RNAs is constitutively strong. As seen in Figure 4, the B-AAV donor, acceptor, and branch point, are more consensus than seen for AAV2, supporting the hypothesis that regulation of AAV2 splicing by its Rep protein depends on its weak intron.

5. AVIAN-ADENO-ASSOCIATED VIRUS (A-AAV)

Avian AAV (A-AAV) was first isolated from the Olson strain of quail bronchitis adenovirus (11). The genome of A-AAV has a genome of 4,694 nucleotides (nt) in length with inverted repeats of 142 nt and has organization similar to that of other AAVs (83). The entire genome of A-AAV displays 56-65% identity at the nucleotide level with the other known AAVs (83), however, the ORF encoding the Rep proteins of A-AAV displays only approximately 50% identity with that of other AAVs.

The intron of A-AAV is larger than that of AAV2. It is 453 nts in length, with a 5' splice site at nt 1824, 3' splice acceptor sites are at nt 2247 (A1) and 2277 (A2), and an internal polyadenylation site p(Ap) (nucleotide numbers refer to GenBank accession number AY186198) (Figure 6). A-AAV RNAs are processed significantly differently than either AAV2 or AAV5. Like AAV5, splicing of A-AAV RNAs shows less dependence in 293 cells on adenovirus. However, in contrast to AAV5, only approximately 50% of the A-AAV RNAs generated from upstream promoters read through (pA)p. A-AAV generates lower levels of spliced P5 and P19 products than does AAV2, suggesting that A-AAV produces low relative levels of Rep 68 and Rep 40. However, expression of the capsid-gene promoter of A-AAV is, like AAV5, largely independent of both adenovirus and its Rep proteins. This

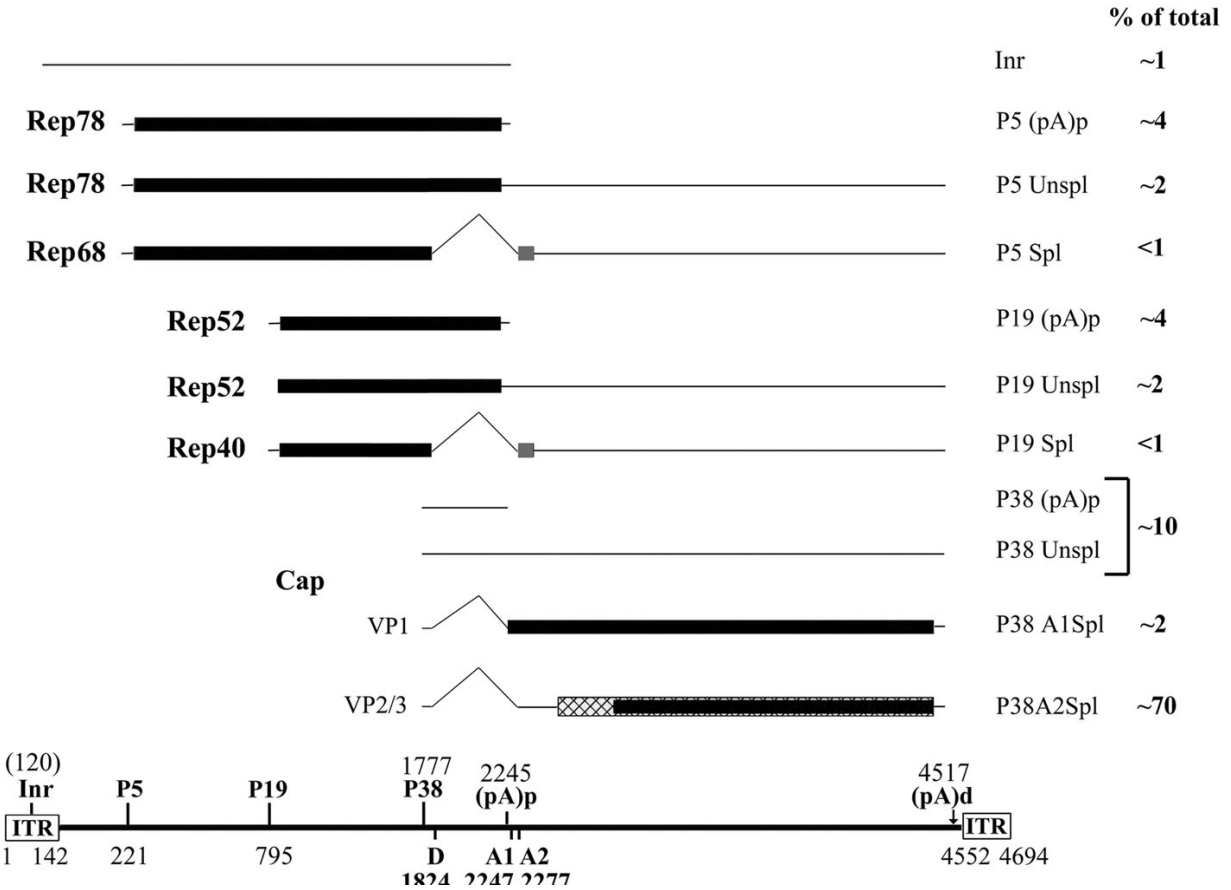

Figure 6. Transcription map of A-AAV. The 4694-nt A-AAV genome (Genbank accession number: AY186198) is shown with transcription landmarks shown to scale, including the ITR, P5 promoter, P19 promoter, P38 promoter, splice donor (D) and acceptor (A1 and A2) sites, the internal polyadenylation site ((pA)p), and the distal polyadenylation site ((pA)d). All of the RNA species are shown with their respective sizes on the right. In contrast to B-AAV and AAV5, a significant percentage of P5- and P19-generated mRNAs read through the internal (pA)p site and are found either as unspliced (unspl) or spliced (spl) species. The approximate relative accumulated levels of each transcript type in total RNA is given as a percentage. All the proteins encoded by A-AAV transcripts are shown with designations on the left.

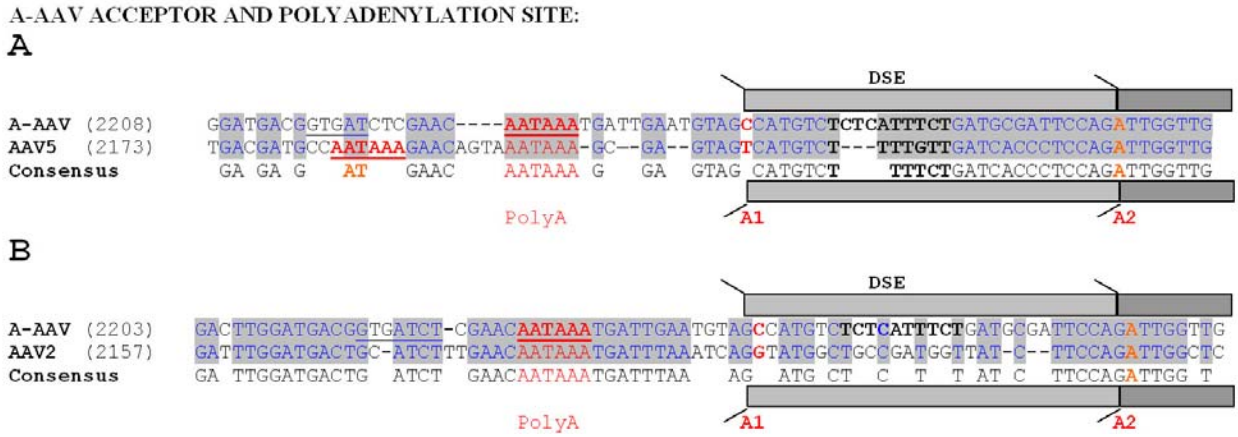

Figure 7. Alignment of the acceptors sites of the A-AAV with AAV5 and AAV2, respectively. The nucleotide sequences of the 3' splice acceptors are aligned for A-AAV and AAV5 (A), or for A-AAV and AAV2 (B). A consensus sequence is presented at the bottom. The intron bordering nucleotides and the polyA signal are shown in red. The used polyA signal is underlined. The DSE region is shown in bold.

suggests that splicing of A-AAV presumably is not regulated by its Rep protein.

Interestingly, although the polypyrimidine region of the A-AAV A2 3' splice site is similar to that of the AAV5, the A-AAV A1 3' splice site is similar to that of AAV2 - only one AAUAAA polyA signal is present, 13 nts in front of the A1 3' cleavage site (20) (Figure 7). In contrast to AAV2, however, this polyA signal is used efficiently (a relative ratio of approximately 50%). One possibility to explain this difference is that the DSE of the A-AAV (pA)p may be stronger than that accompanying AAV2, yet weaker than that present for AAV5 (Figure 7). Thus, the DSE signal seems to be a critical element in the determination of the relative usage of the internal polyA signal. In addition, A-AAV also generates low levels of ITR-initiated transcripts following infection of 293 cells. However, all these transcripts are polyadenylated at (pA)p, which is in contrast to RNAs initiated from the ITRs of AAV5 or B-AAV which uses (pA)p (19,20). The basis for this difference is currently unknown.

6. CONCLUDING REMARKS

AAVs, first identified in the 1960s, have been extensively developed as vectors for human gene therapy in recent years. However, this group of viruses also represents an information dense, tractable system with which to examine eukaryotic RNA processing mechanisms. AAVs are excellent models to understand potential interactions between alternative splicing and alternative polyadenylation, the export of unspliced RNAs, and transactivating proteins that influence RNA processing co-transcriptionally via interaction with the transcription template. Further analysis of AAV RNA processing is sure to shed light on these important cellular processes.

7. ACKNOWLEDGEMENTS

We would like to thank past and present members of our labs for helpful advice and discussion. Work cited from the authors' labs was supported by grants from the National Institutes of Health.

8. REFERENCES

1. R.M.Linden, Molecular biology of adeno-associated viruses., in J.R.Kerr, S.F.Cotmore, M.Bloom, R.M.Linden, and C.R.Parrish (Eds.), Parvoviruses, Hodder Arnold, London, UK, (2006) pp. 68-84.
2. P.Tattersall, The rodent parvoviruses, Viral and mycoplasmas infections of laboratory rodents-Effects on biomedical research., Academic Press, Inc., Florida, (1986) pp. 305-348.
3. P.Tattersall and S.F.Cotmore, Structure and organization of the viral genome., in J.R.Kerr, S.F.Cotmore, M.Bloom, R.M.Linden, and C.R.Parrish (Eds.), Parvoviruses, Hodder Arnold, London, UK, (2006) pp. 73-94.
4. P.Tattersall, The evolution of parvovirus taxonomy., in J.R.Kerr, S.F.Cotmore, M.Bloom, R.M.Linden, and C.R.Parrish (Eds.), Parvoviruses, Hodder Arnold, London, UK, (2006) pp. 5-14.
5. P.Ward, Replication of adeno-associated virus DNA, in J.R.Kerr, S.F.Cotmore, M.Bloom, R.M.Linden, and C.R.Parrish (Eds.), Parvoviruses, Hodder Arnold, London, UK, (2006) pp. 189-212.
6. N.R.Blacklow, M.D.Hoggan, and W.P.Rowe: Serologic evidence for human infection with adenovirus-associated viruses. *J Natl Cancer Inst.* 40 319-327 (1968)
7. B.Georg-Fries, S.Biederlack, J.Wolf, and H.zur Hausen: Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus. *Virology* 134 64-71 (1984)
8. W.P.Parks, J.L.Melnick, R.Rongey, and H.D.Mayor: Physical assay and growth cycle studies of a defective adeno-satellite virus. *J. Virol.* 1 171-180 (1967)
9. U.Bantel-Schaal and H.zur Hausen: Characterization of the DNA of a defective human parvovirus isolated from a genital site. *Virology* 134 52-63 (1984)
10. G.P.Gao, M.R.Alvira, L.Wang, R.Calcedo, J.Johnston, and J.M.Wilson: Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. U. S. A* 99 11854-11859 (2002)
11. V.J.Yates, A.M.el Mishad, K.J.McCormick, and J.J.Trentin: Isolation and characterization of an Avian adenovirus-associated virus. *Infect. Immun.* 7 973-980 (1973)
12. M.F.Coria and H.D.Lehmkuhl: Isolation and identification of a bovine adenovirus type 3 with an adenovirus-associated virus. *Am. J. Vet. Res.* 39 1904-1906 (1978)
13. J.K.Clarke, J.B.McFerran, E.R.McKillop, and W.L.Curran: Isolation of an adeno associated virus from sheep. Brief report. *Arch. Virol.* 60 171-176 (1979)
14. S.L.Farkas, Z.Zadori, M.Benko, S.Essbauer, B.Harrach, and P.Tijssen: A parvovirus isolated from royal python (Python regius) is a member of the genus Dependovirus. *J. Gen. Virol.* 85 555-561 (2004)
15. M.A.Lochrie, G.P.Tatsuno, A.E.Arbetman, K.Jones, C.Pater, P.H.Smith, J.W.McDonnell, S.Z.Zhou, S.Kachi, M.Kachi, P.A.Campochiaro, G.F.Pierce, and P.Colosi: Adeno-associated virus (AAV) capsid genes isolated from rat and mouse liver genomic DNA define two new AAV species distantly related to AAV-5. *Virology.* 353 68-82 (2006)
16. K.Berns and C.R.Parrish, Parvoviridae, in D.M.Knipe and P.M.Howley (Eds.), Fields VIROLOGY, Lippincott Williams & Wilkins, Philadephia, PA, 4th ed., (2006) pp. 2437-2477.
17. B.J.Carter, J.P.Trempe, and E.Mendelson, Adeno-associated virus gene expression and regulation, in P.Tijssen (Ed.), CRC handbook of parvoviruses, CRC Press, Boca Raton,Fla., (1990) pp. 227-254.
18. N.Muzyczka and K.I.Berns, Parvoviridae: The viruses and their replication, in D.M.Knipe (Ed.), Fields' virlogy, Lippincott Williams & Wilkins, Philadelphia, (2001) pp. 2327-2379.
19. J.Qiu, R.Nayak, G.E.Tullis, and D.J.Pintel: Characterization of the transcription profile of AAV5 reveals a number of unique features compared to previously

characterized adeno-associated viruses. *J. Virol.* 76 12435-12447 (2002)
20. J.Qiu, F.Cheng, and D.J.Pintel: Expression profiles of bovine adeno-associated virus and avian adeno-associated virus display significant similarity to that of adeno-associated virus type 5. *J. Virol.* 80 5482-5493 (2006)
21. J.Qiu, F.Cheng, and D.Pintel: Molecular characterization of caprine adeno-associated virus (AAV-Go.1) reveals striking similarity to human AAV5. *Virology.* 356 208-216 (2006)
22. K.I.Berns and C.Giraud: Biology of adeno-associated virus. *Current Topics in Microbiology and Immunology* 218 1-23 (1996)
23. M.W.Myers, C.A.Laughlin, F.T.Jay, and B.J.Carter: Adenovirus helper function for growth of adeno-associated virus: effect of temperature-sensitive mutations in adenovirus early gene region 2. *J. Virol.* 35 65-75 (1980)
24. M.Mouw and D.J.Pintel: Adeno-associated virus RNAs appear in a temporal order and their splicing is stimulated during coinfection with adenovirus. *J. Virol.* 74 9878-9888 (2000)
25. J.P.Trempe and B.J.Carter: Regulation of adeno-associated virus gene expression in 293 cells: control of mRNA abundance and translation. *J. Virol.* 62 68-74 (1988)
26. J.Qiu and D.J.Pintel: The adeno-associated virus type 2 Rep protein regulates RNA processing via interaction with the transcription template. *Mol Cell Biol.* 22 3639-3652 (2002)
27. S.Mori, L.Wang, T.Takeuchi, and T.Kanda: Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein. *Virology.* 320;330 375-383 (2004)
28. E.W.Lusby and K.I.Berns: Mapping of the 5' termini of two adeno-associated virus 2 RNAs in the left half of the genome. *J. Virol.* 41 518-526 (1982)
29. M.R.Green and R.G.Roeder: Definition of a novel promoter for the major adenovirus-associated virus mRNA. *Cell* 22 231-242 (1980)
30. J.A.King, R.Dubielzig, D.Grimm, and J.A.Kleinschmidt: DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J.* 20 3282-3291 (2001)
31. T.H.Ni, X.Zhou, D.M.McCarty, I.Zolotukhin, and N.Muzyczka: In vitro replication of adeno-associated virus DNA. *J. Virol.* 68 1128-1138 (1994)
32. J.P.Trempe and B.J.Carter: Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein. *J. Virol.* 62 3356-3363 (1988)
33. S.P.Becerra, F.Koczot, P.Fabisch, and J.A.Rose: Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript. *J. Virol.* 62 2745-2754 (1988)
34. T.R.Flotte, S.A.Afione, R.Solow, M.L.Drumm, D.Markakis, W.B.Guggino, P.L.Zeitlin, and B.J.Carter: Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. *J. Biol. Chem.* 268 3781-3790 (1993)
35. T.R.Flotte, R.Solow, R.A.Owens, S.Afione, P.L.Zeitlin, and B.J.Carter: Gene expression from adeno-associated virus vectors in airway epithelial cells. *Am. J. Respir. Cell Mol. Biol.* 7 349-356 (1992)
36. R.P.Haberman, T.J.McCown, and R.J.Samulski: Novel transcriptional regulatory signals in the adeno-associated virus terminal repeat A/D junction element. *J. Virol.* 74 8732-8739 (2000)
37. M.A.Labow, P.L.Hermonat, and K.I.Berns: Positive and negative autoregulation of the adeno-associated virus type 2 genome. *J. Virol.* 60 251-258 (1986)
38. L.Mishra and J.A.Rose: Adeno-associated virus DNA replication is induced by genes that are essential for HSV-1 DNA synthesis. *Virology* 179 632-639 (1990)
39. R.M.Buller, J.E.Janik, E.D.Sebring, and J.A.Rose: Herpes simplex virus types 1 and 2 completely help adenovirus-associated virus replication. *J. Virol.* 40 241-247 (1981)
40. B.J.Carter, Adeno-associated virus helper functions, in P.Tijssen (Ed.), CRC handbook of parvoviruses, CRC Press, Boca Raton, Fla., (1990) pp. 255-282.
41. K.Berns, *Parvoviridae*: The Viruses and Their Replication, in B.N.Fields, D.M.Knipe, and P.M.Howley (Eds.), Fundamental Virology, Lippincott-Raven, (1996) pp. 1017-1041.
42. F.K.Ferrari, T.Samulski, T.Shenk, and R.J.Samulski: Second-strand synthesis is a rate limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J. Virol.* 70 3227-3234 (1996)
43. E.Seto, Y.Shi, and T.Shenk: YY1 is an initiator sequence-binding protein that directs and activates transcription *in vitro*. *Nature* 354 241-245 (1991)
44. R.J.Samulski and T.Shenk: Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J Virol.* 62 206-210 (1988)
45. J.E.Janik, M.M.Huston, and J.A.Rose: Locations of adenovirus genes required for the replication of adenovirus-associated virus. *Proc. Natl. Acad. Sci. U. S. A* 78 1925-1929 (1981)
46. P.Ward, F.B.Dean, M.E.O'Donnell, and K.I.Berns: Role of the adenovirus DNA-binding protein in in vitro adeno-associated virus DNA replication. *J. Virol.* 72 420-427 (1998)
47. W.D.Richardson and H.Westphal: A cascade of adenovirus early functions is required for expression of adeno-associated virus. *Cell* 27 133-141 (1981)
48. X.Xiao, J.Li, and R.J.Samulski: Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J. Virol.* 72 2224-2232 (1998)
49. F.W.Weindler and R.Heilbronn: A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. *J. Virol.* 65 2476-2482 (1991)
50. D.L.Bentley: Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* 17 251-256 (2005)
51. T.H.Stracker, G.D.Cassell, P.Ward, Y.M.Loo, B.van Breukelen, S.D.Carrington-Lawrence, R.K.Hamatake, P.C.van der Vilet, S.K.Weller, T.Melendy, and M.D.Weitzman: The Rep protein of adeno-associated virus type 2 interacts with single-stranded DNA-binding proteins that enhance viral replication. *J. Virol.* 78 441-453 (2004)
52. E.Bridge, M.Carmo-Fonseca, A.Lamond, and U.Pettersson: Nuclear organization of splicing small nuclear ribonucleoproteins in adenovirus-infected cells. *J. Virol.* 67 5792-5802 (1993)
53. E.Bridge, K.U.Riedel, B.M.Johansson, and U.Pettersson: Spliced exons of adenovirus late RNAs

colocalize with snRNP in a specific nuclear domain. *J. Cell Biol.* 135 303-314 (1996)
54. A.Pombo, J.Ferreira, E.Bridge, and M.Carmo-Fonseca: Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells. *EMBO J.* 13 5075-5085 (1994)
55. D.M.McCarty, D.J.Pereira, I.Zolotukhin, X.Zhou, J.H.Ryan, and N.Muzyczka: Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. *J. Virol.* 68 4988-4997 (1994)
56. B.A.Antoni, A.B.Rabson, I.L.Miller, J.P.Trempe, N.Chejanovsky, and B.J.Carter: Adeno-associated virus Rep protein inhibits human immunodeficiency virus type 1 production in human cells. *J. Virol.* 65 396-404 (1991)
57. A.R.Kornblihtt: Promoter usage and alternative splicing. *Curr. Opin. Cell Biol.* 17 262-268 (2005)
58. C.N.Cole and J.J.Scarcelli: Transport of messenger RNA from the nucleus to the cytoplasm. *Curr. Opin. Cell Biol.* 18 299-306 (2006)
59. M.H.Malim, J.Hauber, S.Y.Le, J.V.Maizel, and B.R.Cullen: The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* 338 254-257 (1989)
60. L.E.Babiss, H.S.Ginsberg, and J.E.Darnell, Jr.: Adenovirus E1B proteins are required for accumulation of late viral mRNA and for effects on cellular mRNA translation and transport. *Mol. Cell Biol.* 5 2552-2558 (1985)
61. D.N.Halbert, J.R.Cutt, and T.Shenk: Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *J. Virol.* 56 250-257 (1985)
62. K.N.Leppard and T.Shenk: The adenovirus E1B 55 kd protein influences mRNA transport via an intranuclear effect on RNA metabolism. *EMBO J.* 8 2329-2336 (1989)
63. S.Pilder, M.Moore, J.Logan, and T.Shenk: The adenovirus E1B-55K transforming polypeptide modulates transport or cytoplasmic stabilization of viral and host cell mRNAs. *Mol. Cell Biol.* 6 470-476 (1986)
64. E.Bridge and G.Ketner: Redundant control of adenovirus late gene expression by early region 4. *J. Virol.* 63 631-638 (1989)
65. J.Qiu, R.Nayak, and D.J.Pintel: Alternative polyadenylation of adeno-associated virus type 5 RNA within an internal intron is governed by a downstream element within the intron 3' splice acceptor and an element upstream of the P41 initiation site. *J. Virol.* 78 83-93 (2004)
66. R.Javahery, A.Khachi, K.Lo, B.Zenzie-Gregory, and S.T.Smale: DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol. Cell Biol.* 14 116-127 (1994)
67. C.Ye, J.Qiu, and D.J.Pintel: Efficient expression of the adeno-associated virus type 5 p41 capsid gene promoter in 293 cells does not require Rep. *J. Virol.* 80 6559-6567 (2006)
68. J.Zhao, L.Hyman, and C.Moore: Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. and Molecular Biol. Rev.* 63 405-445 (1999)
69. J.Qiu and D.J.Pintel: Alternative polyadenylation of adeno-associated virus type 5 RNA within an internal intron is governed by the distance between the promoter and the intron and is inhibited by U1 small nuclear RNP binding to the intervening donor. *J. Biol. Chem.* 279 14889-14898 (2004)
70. J.Qiu, F.Cheng, and D.Pintel: Distance-dependent processing of adeno-associated virus type 5 RNA is controlled by 5' exon definition. *J. Virol.* 81 7974-7984 (2007)
71. J.D.Lewis, E.Izaurralde, A.Jarmolowski, C.Meguigan, and I.W.Mattaj: A nuclear cap-binding complex facilitates association of U1-snRNP with the cap-proximal 5' splice-site. *Genes Dev.* 10 1683-1698 (1996)
72. E.Izaurralde, J.Lewis, C.McGuigan, M.Jankowska, E.Darzynkiewicz, and I.W.Mattaj: A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* 78 657-668 (1994)
73. E.Izaurralde, J.Lewis, C.Gamberi, A.Jarmolowski, C.McGuigan, and I.W.Mattaj: A cap-binding protein complex mediating U snRNA export. *Nature* 376 709-712 (1995)
74. S.M.Flaherty, P.Fortes, E.Izaurralde, I.W.Mattaj, and G.M.Gilmartin: Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc. Natl. Acad. Sci. U. S. A.* 94 11893-11898 (1997)
75. K.Inoue, M.Ohno, H.Sakamoto, and Y.Shimura: Effect of the cap structure on pre-mRNA splicing in Xenopus oocyte nuclei. *Genes Dev.* 3 1472-1479 (1989)
76. H.V.Colot, F.Stutz, and M.Rosbash: The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev.* 10 1699-1708 (1996)
77. P.A.Furth, W.T.Choe, J.H.Rex, J.C.Byrne, and C.C.Baker: Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol. Cell Biol.* 14 5278-5289 (1994)
78. S.I.Gunderson, M.Polycarpou-Schwarz, and I.W.Mattaj: U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell.* 1 255-264 (1998)
79. P.Fortes, Y.Cuevas, F.Guan, P.Liu, S.Pentlicky, S.P.Jung, M.I.Martinez-Chantar, J.Prieto, D.Rowe, and S.I.Gunderson: Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small nuclear RNAs targeted to terminal exons of pre-mRNA. *Proc. Natl. Acad. Sci. USA* 100 8264-8269 (2003)
80. M.Schmidt, H.Katano, I.Bossis, and J.A.Chiorini: Cloning and characterization of a bovine adeno-associated virus. *J. Virol.* 78 6509-6516 (2004)
81. E.J.Olson, S.R.Haskell, R.K.Frank, H.D.Lehmkuhl, L.A.Hobbs, J.V.Warg, J.G.Landgraf, and A.Wunschmann: Isolation of an adenovirus and an adeno-associated virus from goat kids with enteritis. *J. Vet. Diagn. Invest.* 16 461-464 (2004)
82. A.E.Arbetman, M.Lochrie, S.Zhou, J.Wellman, C.Scallan, M.M.Doroudchi, B.Randlev, S.Patarroyo-White, T.Liu, P.Smith, H.Lehmkuhl, L.A.Hobbs, G.F.Pierce, and P.Colosi: Novel caprine adeno-associated virus (AAV) capsid (AAV-Go.1) is closely related to the primate AAV-5 and has unique tropism and neutralization properties. *J. Virol.* 79 15238-15245 (2005)

83. I.Bossis and J.A.Chiorini: Cloning of an avian adeno-associated virus (AAAV) and generation of recombinant AAAV particles. *J. Virol.* 77 6799-6810 (2003)

Key Words: Adeno-Associated Viruses, Parvoviruses, RNA Processing, Transcription, Review

Send correspondence to: Dr David Pintel, 471f Life Sciences Center, University of Missouri-Columbia, 1201 Rollins Rd, Columbia, MO 65212, Tel: 573-0882-3920, Fax: 573-884-9676, E-mail: pinteld@missouri.edu

http://www.bioscience.org/current/vol13.htm