JOURNAL OF VIROLOGY, Sept. 2000, p. 8732–8739 Vol. 74, No. 18
0022-538X/00/$04.00+0
Copyright © 2000, American Society for Microbiology. All Rights Reserved.

Novel Transcriptional Regulatory Signals in the Adeno-Associated Virus Terminal Repeat *A/D* Junction Element

REBECCA P. HABERMAN,[1,2] THOMAS J. McCOWN,[1,3,4] AND RICHARD JUDE SAMULSKI[1,2,5]*

UNC Gene Therapy Center,[1] Neuroscience Center,[3] Curriculum in Neurobiology,[2] and Departments of Psychiatry[4] and Pharmacology,[5] University of North Carolina, Chapel Hill, North Carolina 27599

Received 7 March 2000/Accepted 3 June 2000

Adeno-associated virus (AAV) type 2 vectors transfer stable, long-term gene expression to diverse cell types in vivo. Many gene therapy applications require the control of long-term transgene expression, and AAV vectors, similar to other gene transfer systems, are being evaluated for delivery of regulated gene expression cassettes. Previously, we (R. P. Haberman, T. J. McCown, and R. J. Samulski, Gene Ther. 5:1604–1611, 1998) demonstrated the use of the tetracycline-responsive system for long-term regulated expression in rat brains. In that study, we also observed residual expression in the "off" state both in vitro and in vivo, suggesting that the human cytomegalovirus (CMV) major immediate-early minimal promoter or other *cis*-acting elements (AAV terminal repeats [TR]) were contributing to this activity. In the present study, we identify that the AAV TR, minus the tetracycline-responsive minimal CMV promoter, will initiate mRNA expression from vector templates. Using deletion analysis and specific PCR-derived TR reporter gene templates, we mapped this activity to a 37-nucleotide stretch in the *A/D* elements of the TR. Although the mRNA derived from the TR is generated from a non-TATA box element, the use of mutant templates failed to identify function of canonical initiator sequences as previously described. Finally, we demonstrated the presence of green fluorescent protein expression both in vitro and in vivo in brain by using recombinant virus carrying only the TR element. Since the AAV terminal repeat is a necessary component of all recombinant AAV vectors, this TR transcriptional activity may interfere with all regulated expression cassettes and may be a problem in the development of novel TR split gene vectors currently being considered for genes too large to be packaged.

Recombinant adeno-associated virus (AAV) vectors have been studied in both large- and small-animal models and demonstrate safe and effective gene transfer (22). Importantly, AAV vectors transduce tissues for long periods, lasting over 1.5 years in the muscle and central nervous system of rodents (27, 37). Such long-term expression suggests the possibility of permanent gene transfer in human gene therapy. While this is essential for the correction of genetic diseases, it raises the new concern of properly regulating the therapeutic gene product.

Appropriate implementation of long-term gene expression requires the ability to control the expression of virus-delivered transgenes using either endogenous cell-type-specific promoters or exogenous regulation systems. One exogenous system, the tetracycline-regulated system, has been studied extensively in vitro and in vivo and has been shown to give 100- to 1,000-fold levels of regulation between the off and on states (11, 15). By incorporating this system in an AAV vector, we demonstrated control of reporter gene expression in the brain (13), with a range of regulation of 28-fold in vitro and 10-fold in vivo. Two other groups have generated tetracycline-regulated AAV vectors that demonstrated control of the secreted erythropoietin protein after vector injection into rodent muscle, with similar success (1, 18). While the ability to obtain regulated gene expression in these studies was significant, the results were vastly different from those published using transfections in vitro or analysis of tetracycline regulation in transgenic animals (100- to 1,000-fold). It is clear that many human gene therapy and experimental situations will require tighter control. Our data suggest that the decreased regulation is most likely to be due to a high background level of activity in the "off" state. Since AAV vectors are derived from minimal *cis*-acting sequences (145-bp terminal repeats), only a limited number of components of the vector could be responsible for this activity. This includes a portion of the tetracycline-regulated promoter, the cytomegalovirus (CMV) minimal promoter, and/or the AAV terminal repeats (TR). Background expression from the minimal promoter can be significant, and new versions of the tetracycline-regulated system are showing promise in resolving this concern (10). However, several lines of evidence suggest that the AAV TR may also be a source of transcriptional elements (8, 9) that could adversely influence regulation from the tetracycline-regulated AAV vector.

Since the AAV TRs contain all the *cis*-acting sequences necessary for replication and packaging of recombinant DNA as well as for mediating the integration of the viral DNA into the host genome, they cannot be deleted. These sequences have at least two different transcriptional activities in tissue culture (8, 9). Flotte et al. (9) removed the first 83 or 140 nucleotides (nt) of the AAV TR located adjacent to a TATA box-containing promoter and reduced expression by about 50% after plasmid transfections into an epithelial cell line. This suggested that an enhancer activity was located somewhere in the first 83 nt of the AAV TR. In a second study (8), they demonstrated that the AAV TR was able to initiate gene expression in a promoterless construct after plasmid transfection or after AAV vector infection in vitro. Thus, the TR appears to have both promoter and enhancer activities in tissue culture. While these studies identified enhancer and transcriptional activities, further identification of the critical sequences responsible for these observations have not been forthcoming, probably due to the difficulty of characterizing the AAV TR.

The TR contains significant secondary structure (see Fig. 2A), which may also be responsible for the reported transcrip-

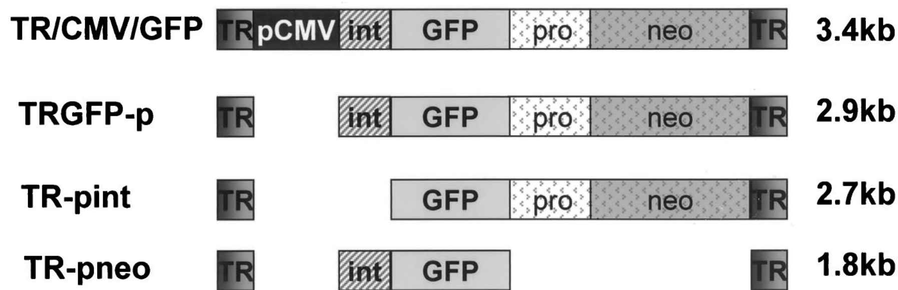
* Corresponding author. Mailing address: CB 7352, 7119 Thurston Bowles, UNC Gene Therapy Center, University of North Carolina, Chapel Hill, NC 27599. Phone: (919) 962-3285. Fax: (919) 966-0907. E-mail: rjs@med.unc.edu.

tional activity. Due to palindromic sequences within the TR, it is the only double-stranded portion of an otherwise single-stranded AAV genome when packaged (7). This sequence, which has a G+C content of over 80%, folds back on itself to form a T-shaped structure composed of two small palindromes, *B* and *C*, contained within a larger palindrome, *A* (see Fig. 2A). An additional sequence, the *D* element, is also part of the TR but is single stranded in the virion. The *D* element contains the packaging signal (X. Xiao and R. J. Samulski, unpublished data), and the *A* stem contains sequences for an AAV Rep protein binding element (RBE), while the junction between these two elements defines the terminal resolution site (*trs*) required for viral DNA replication (2, 14, 21, 28). In addition to the general secondary structure which may attract transcriptional proteins (30, 34) and as initially proposed by Flotte et al. (8), we identified two sequences in the *A* (nt 88 to 95) and *D* (nt 126 to 133) regions which have significant homology (5 of 7 bases for each) to an initiator sequence (see Fig. 2A). Initiator sequences are characterized as minimal sequences that are sufficient to mediate transcript initiation in the absence of a TATA box (25, 26). Flotte et al. (8) suggested that a putative initiator sequence may generate the promoter activity seen in their studies with AAV. It is possible that these sequences in the TR are sufficient to produce transcriptional activation of a downstream gene in vitro and in vivo. To better characterize these observations, we initiated studies aimed at defining the role of the TR on the tetracycline-responsive vectors we used in vivo.

TR-initiated expression. To better define the reason for inefficient regulation of our tetracycline-responsive vector, a series of promoterless plasmids were derived from TR/CMV/GFP by restriction enzyme digestion (Fig. 1A). TRGFP-p is identical to TR/CMV/GFP except that it is missing the CMV promoter that is typically used to express the reporter gene, green fluorescent protein (GFP). Superfect transfection (1 μg of plasmid to 4 μl of Superfect reagent [Qiagen]) of TRGFP-p into approximately 2×10^5 HEK 293 cells (data not shown) or 1×10^5 HtTA-1 cells (HeLa cells stably transfected with the tetracycline transactivator [a gift from H. Bujard]) expressed GFP at a low level (Fig. 1B). In HtTA-1 cells, TRGFP-p produced detectable expression in 5 to 10% of the cells that showed expression with TR/CMV/GFP and at a lower intensity per cell. To ensure that this activity was not due to other elements (e.g., the simian virus 40 [SV40] intron sequence or plasmid readthrough by the neomycin gene promoter), TR-pint and TR-pneo were constructed to remove these sequences, respectively. Transfection of each construct produced gene expression at a level equal to or slightly greater than that for TRGFP-p. Thus, the AAV TR in plasmid form (Fig. 2) appears sufficient to produce low-level transcriptional activity in tissue culture. These observations corroborated the published results of Flotte et al. (8) using chloramphenicol acetyltransferase reporter gene in 293 cells.

Mapping TR promoter activity. Localization of the promoter activity within the TR was investigated initially by primer extension analysis of the mRNA start site. Our technique produced clear identification of a CMV promoter start site but inconclusive results with the TR-generated transcripts, presumably due to low levels of mRNA. Therefore, reverse transcription-PCR (RT-PCR) analysis of mRNA from plasmid substrates transfected into 293 cells was initiated to determine an approximate start site. Products of mRNA amplification were distinguished from any input DNA amplification by use of an intron-containing construct. The DNA PCR products also served as an internal control for the reaction (since they were 100 bp larger than the spliced mRNA product). A series of primers (Fig. 2) were used to amplify products of spliced mRNA transcripts and 100-bp-larger products from unspliced RNA or residual DNA (TR-pneo). Primers located 5′ of the initiation site would produce products only of unspliced length, indicating amplification of residual DNA and not processed mRNA. mRNA was extracted using Trizol reagent (Gibco BRL) from 10-cm plates of TR-pneo-transfected 293 cells and Dynal poly-T magnetic beads (as specified by the manufacturer). The mRNA was then subjected to RT with poly(T) primer and avian myeloblastosis virus reverse transcriptase for 90 min at 42°C. The RT products were PCR amplified with a series of 5′ TR primers and one of two 3′ primers located in the GFP gene, under conditions optimized to detect low transcript numbers (1 U of Perkin-Elmer *Taq* polymerase, 0.2 μM each primer, and 0.25 mM each deoxynucleoside triphosphate per 50-μl reaction mixture with 40 cycles of amplification). Gel electrophoresis analysis of the PCR products revealed both spliced and unspliced products from primers 2 through 6 (Fig. 2B, lanes 1, 4, and 7; Table 1), while mRNA preparations from mock-transfected cells, spiked with DNA, yielded only unspliced sized products for these primers. Although the TR has many repetitive sequences, no internal TR amplification (indicated by smaller bands) was seen with these primers. Primer 7, located in the B/C region of the terminal repeat, was unable to amplify any products of the spliced or unspliced size, even with the use of a series of annealing temperatures and a GC-rich sequence optimization reagent (GC melt; Clontech). This primer is located in a GC-rich region in the middle of the secondary structure of the terminal repeat (Fig. 2A, virus form). Previous studies have determined that this secondary structure interferes with PCR amplification (23, 38) and is probably blocking amplification from primer 7. Likewise, primers located beyond the terminal repeat in the plasmid were unable to consistently amplify products of either spliced or unspliced size, confirming previous observations (28, 38) and making the additional localization of transcript initiation beyond this secondary structure intractable. To ensure that the spliced products detected with the initial primers were amplified from RNA, mRNA preparations were treated with 30 U of S1 nuclease or 0.3 M NaOH for 1 h at 37°C in the presence of glycogen and then ethanol precipitated before being subjected to the RT step. Since S1 nuclease digestion and NaOH digestion preferentially degrade single-stranded nucleic acids and RNA, respectively, only products derived from DNA (unspliced size) were detected after these treatments (Fig. 2B, lanes 2 and 3; Table 1). This approach confirmed that mRNA transcripts originated as far back as the 5′ end of the *A* element (primer 6). However, because of our inability to amplify any product upstream of primer 6, additional determinations of any other 5′ initiation sequences were not possible using this method.

To further evaluate the TR promoter activity and the minimal sequences necessary for mRNA expression, PCR amplification was used to generate minimal TR/GFP cassettes, which were purified, transfected into HtTA-1 or 293 cells, and assayed for GFP activity. A number of the primers used to map mRNA initiation from plasmid vector templates, along with several additional 5′ primers paired with a 3′ primer located after the poly(A) signal of the GFP gene, were used to generate these miniexpression cassettes (Fig. 3A). PCR products were amplified from plasmid templates (TR-pneo or TR-pint) under conditions similar to those for RT product amplification. These products were amplified a second time to obtain enough material for transfection into duplicate wells of a 12-well plate. After PCR, the products were digested with *Dpn*I, which digests input plasmid DNA, for 2 h and purified on Sephadex



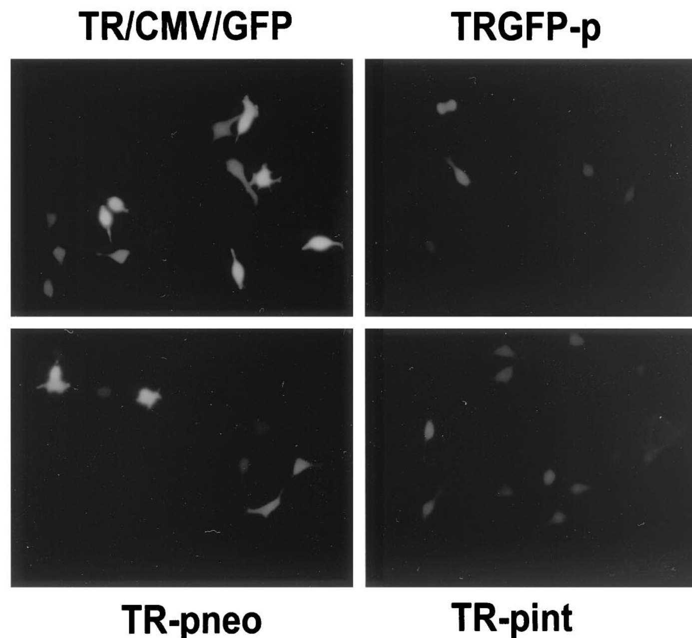
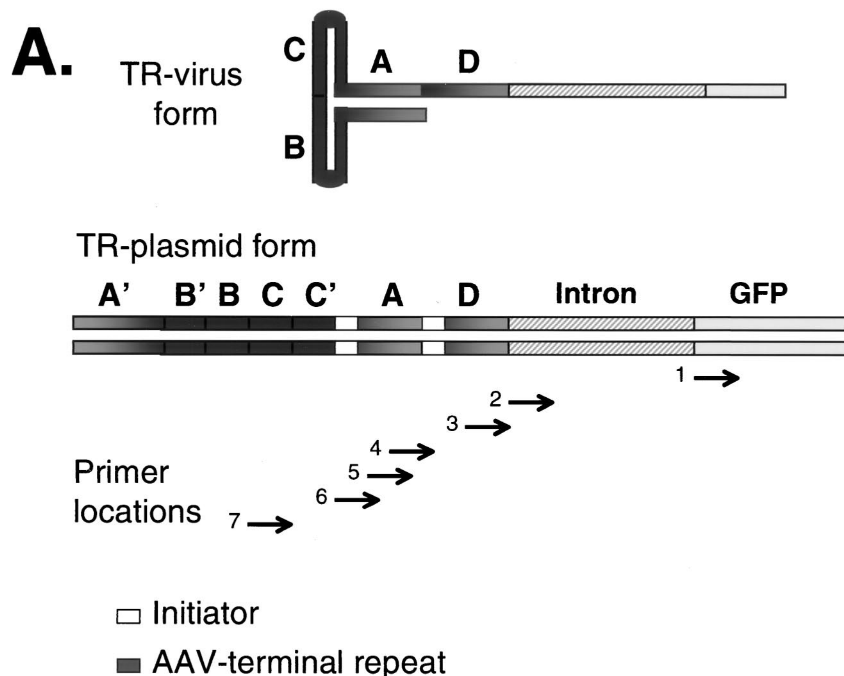


FIG. 1. (A) Plasmids used for transfection. All promoterless plasmids are derived from TR-CMV-GFP by restriction digest deletion of the regions indicated. TR-CMV-GFP was made from pTRUF 2 (39), a kind gift from N. Muzyczka, by removing the GFP gene and replacing it with the EGFP gene (Clonetech) using convenient restriction enzymes. pCMV, cytomegalovirus immediate-early promoter; int, SV40 intron; pro, TK promoter plus tandem repeats of a polyomavirus enhancer; neo, neomycin resistance gene. (B) Photomicrographs of HtTA-1 cells 24 h after transfection of the indicated plasmid. Exposure times are not equivalent for all photomicrographs.

G-25 spin columns to remove residual deoxynucleoside triphosphates and primers. Products were quantified by optical densitometry, and approximately 1.5 pmol of each product was transfected into 293 cells or HtTA-1 cells using superfect (in quantification experiments, equivalent amounts [in picomoles] of each product were used). Two measurements were used to determine relative expression: the number of GFP-positive cells and relative whole-field fluorescence (Fig. 3A). Templates containing only the intron (amplified using primer 2) and no TR sequences produced a few visible cells (the number of GFP-positive cells was 18.9% ± 5.9% of the number obtained with control primer 5) after transfection (Fig. 3). This expression could be due to nonspecific or random initiation within the SV40 intron. However, the levels significantly increased (the number of GFP-positive cells was 91.8% ± 9.1% of the number with control primer 5) after transfection of a PCR template that included sequences of the TR, specifically the *A*/*D* junction (primer 4). Since intron-minus templates generated from primer 4 also gave positive expression (Fig. 3B), we concluded that the intron was not required for this increase. PCR templates carrying a small portion of the *D* element (primer 3) produced slightly lower expression than that obtained with primer 2 (intron primer). This observation was very reproducible, indicating the ability of the intron sequences to initiate nonspecifically or the potential interference from the cellular protein(s) previously shown to bind the AAV *D* ele-



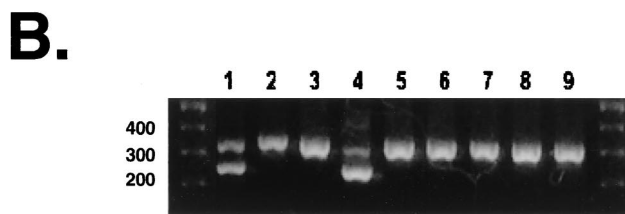


FIG. 2. (A) Diagram of the TR in genome form as packaged in the virion (top) and as cloned into a plasmid (bottom). The TR forms a T-shaped structure containing one continuous strand in the virion. In the plasmid the TR is double stranded and contains the *A*, *B*, and *C* regions and their complements, *A*′, *B*′, and *C*′. The location within the terminal repeat, intron, or GFP gene of 5′ primers 1 to 7 are also shown. Primer 1 spans −13 to +6 of the GFP gene, primer 2 spans −3 to +15 of the SV40 intron (sequences 5′ of the GFP gene and the intron are polylinker regions), primer 3 spans nt 139 to 145 of the TR plus 15 nt of the polylinker, primer 4 spans nt 110 to 127 of the TR, primer 5 spans nt 104 to 122, primer 6 spans nt 88 to 105, and primer 7 spans nt 60 to 77. (B) RT-PCR products were amplified using primer 2 from RNA isolated from transfected 293 cells and analyzed by agarose gel electrophoresis. Lanes: 1 to 3, TR-pneo (Fig. 1)-transfected cells; 4 to 6, TR/CMV/GFP-transfected cells; 7 to 9, mock-transfected cells that were spiked with TR-pneo DNA at cell harvest; 1, 4, and 7, untreated mRNA; 2, 5, and 8, RNA samples treated with S1 nuclease prior to RT-PCR; 3, 6, and 9, RNA samples treated with NaOH to degrade RNA prior to RT-PCR. The 3′ primer paired with primer 2 was located in the GFP gene: GFP-RT2 (nt 145 to 126 of the EGFP gene). Either GFP-RT2 or GFP-RT1 (nt 19 to 3 of the EGFP gene) was paired with the 5′ primers from panel A, as delineated in Table 1.

ment (17). As an additional control, PCR templates generated from a primer containing the GFP start codon (primer 1) gave no expression (Fig. 3A), supporting the premise that transcriptional activation observed in this assay was due primarily to upstream sequences in the TR. Since templates derived from primer 4 included the putative *D* element initiator sequence, two additional variants of primer 4 carrying severe mutations of the canonical initiator sequence as previously described were tested (25, 26). PCR products generated from these mutations had no consistent effect on GFP expression, indicating that the initiator sequence, as currently identified, is not responsible for TR transcript initiation observed in our assays. Likewise, hybrid templates, generated using primer 8, which carries the *A* element initiator-like sequence coupled to the insufficient primer 3 sequence amplified from an *A* element-deleted plasmid (TR-pneo/mp), did not increase GFP expression beyond that observed with primer 3 alone. This suggests

TABLE 1. Primers used in this study

5′ primer[a]	3′ primer	Unspliced product detected[b]	Spliced product detected[b]	Spliced product NaOH and S1 sensitive
Primer 2	GFP-RT2	+	+	Yes
Primer 3	GFP-RT2	+	+	Yes
Primer 4	GFP-RT1	+	+	NT[c]
Primer 5	GFP-RT1	+	+	NT
Primer 5	GFP-RT2	+	+	NT
Primer 6	GFP-RT1	+	+	NT
Primer 6	GFP-RT2	+	+	NT
Primer 7	GFP-RT1	−	−	NT

[a] A series of 5′ primers were used with RT-PCR products from RNA isolated from TR-pneo-transfected cells (Fig. 2). Primer sequences and locations are described in the legend to Fig. 2.

[b] +, product of the correct size detected; −, no product of the correct size detected.

[c] NT, not tested.

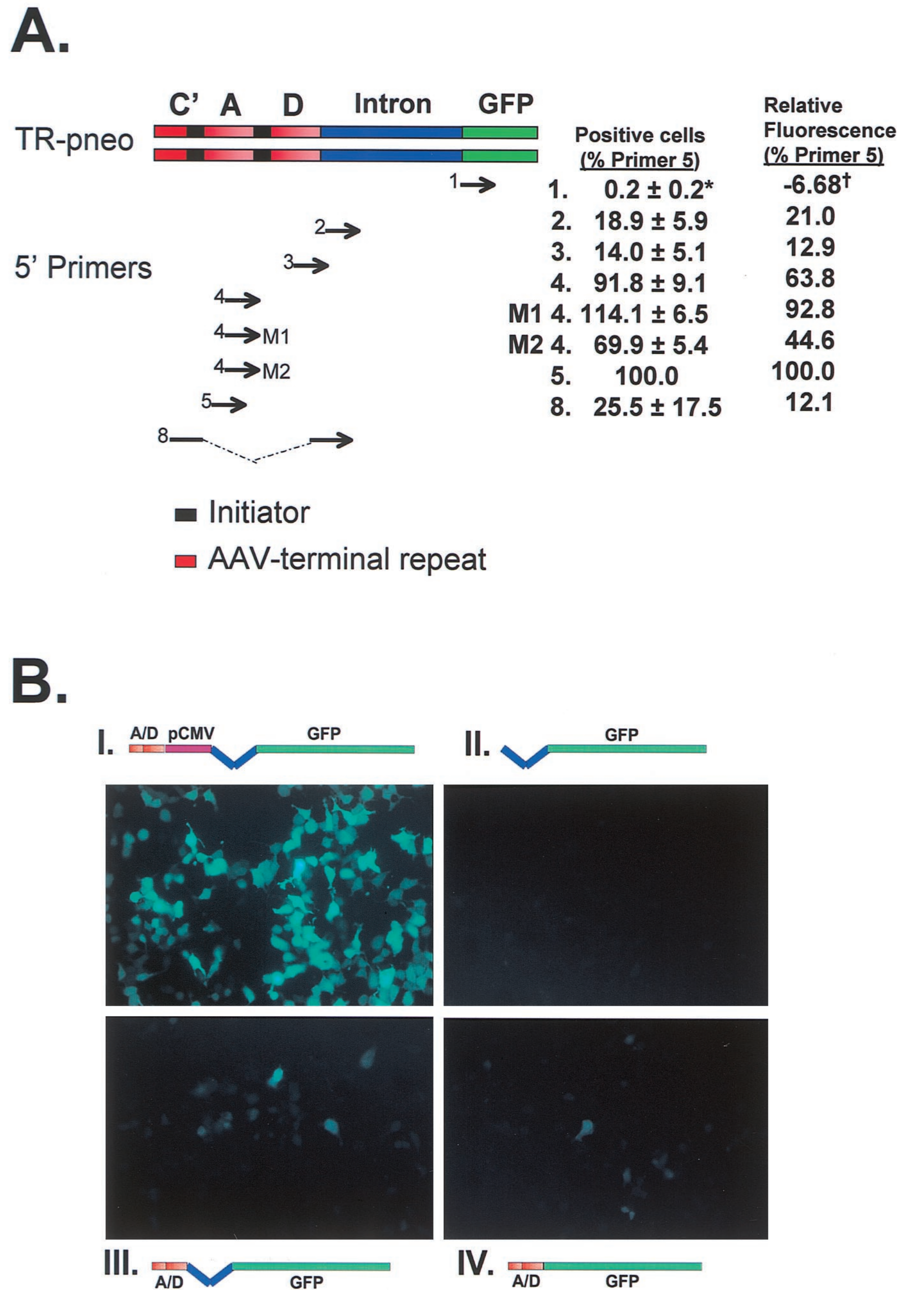
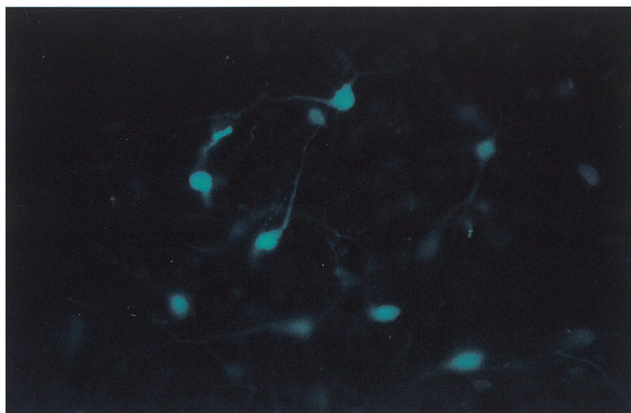


FIG. 3. (A) 5′ PCR primers to the indicated regions were paired with a GFP poly(A) region primer to amplify products containing portions of the TR or intron and the entire GFP gene. After transfection of the PCR products into HtTA-1 cells, relative positive cell counts and relative field fluorescence were determined from a random field with a 10× objective (eight fields per well, two wells per condition). Cell counts were obtained for three independent experiments and are expressed as a mean percentage of that obtained for primer 5 ± the standard error. Whole-field fluorescence, averaged from two of the same experiments, was corrected for cell autofluorescence by subtracting the background of mock-transfected cells and are also expressed as a percentage of the intensity obtained with primer 5. Primer 4 M1 contains nt 110 to 138 of the TR, except that CCA at nt 128 to 130 was changed to GGG. Primer 4 M2 contains nt 110 to 138 of the TR, except that CT at nt 126 and 127 was changed to GA. Primer 8 consists of nt 72 to 93 linked to nt 139 to 145 of the TR plus 15 nt of polylinker (primer 3) and amplified from TR-pneo/mp. TR-pneo/mp is derived from TR-pneo by digestion with *Msc*I and *Pst*I, blunting, and then religating the backbone with the GFP fragment. This plasmid contains only nt 119 to 145 of the TR. *, two of three experiments with this primer showed no positive cells (in one experiment, two faint green cells of unknown origin were detected); †, because whole-field measurements cannot distinguish between specific GFP fluorescence and cell autofluorescence due to cell death, these values contain more variability than the cell count measurement. (B) Photomicrographs of 293 cells 48 h after transfection of the PCR products illustrated above each photograph. Exposure times were equivalent. All PCR products were derived using the same 3′ poly(A) signal primer. Plasmid and 5′ primers were as follows: I, TR/CMV/GFP (Fig. 1) with primer 4; II, TR-pneo (Fig. 1) with primer 2; III, TR-pneo with primer 4; IV, TR-pint (Fig. 1) with primer 4.

that the putative *A* element initiator is not sufficient to initiate transcription under these conditions.

We determined that sequences of the TR amplified by primer 4 (nt 109 to 145) are sufficient to express GFP from a linear DNA template in a transfection assay. To ensure that these results were not related to the total length of the 5′ sequences, our PCR products generated with primer 8 (Fig. 3A) contained the same amount of 5′ DNA as did templates generated with



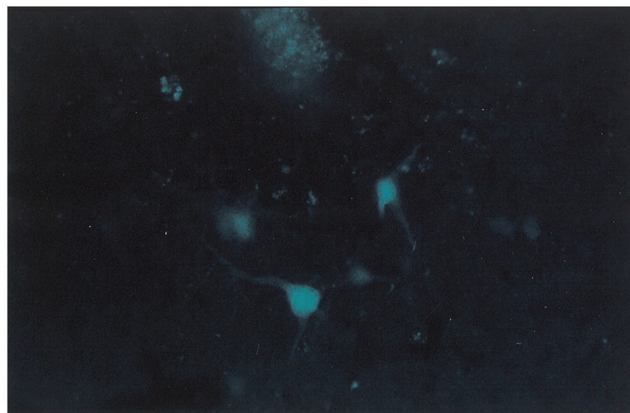


FIG. 4. In vivo expression of TR-driven GFP. Virus made from a TR-pneo (AAV TR-pneo) was injected into the inferior colliculus and the hypothalamus of rats; 2 weeks later, the brains were sectioned and analyzed for GFP fluorescence by fluorescent microscopy. Many cells from each animal are visible in both regions of the brain.

primer 4. Even under these conditions, primer 8-derived templates did not yield efficient GFP expression. Although the TR sequence appeared to contain a classic initiator sequence, as determined by sequence homology of published initiator elements, results obtained with PCR templates generated from specific mutant initiator primers strongly suggest otherwise. Our study shows that the AAV-TR element has RNA initiation capacity, maps to nt 109 to 145, and does not use a classic initiator-like element to generate this activity. No other obvious sequence motif could be identified as a candidate for mRNA initiation; therefore, further experiments are required to determine the exact mechanism of mRNA initiation from this region. While our experiments clearly demonstrate a role for this sequence in the PCR-based transfection assay, we extended these observations by testing for GFP expression by generating a promoter-minus recombinant AAV and assaying in vivo.

AAV-TR vector expression in vivo. It is clear from the experiments above that in plasmid or PCR-derived linear DNA form, the AAV terminal repeat can initiate the transcription of a reporter gene (the PCR template gives 2.9% of the expression obtained with the CMV template). However, this expression is only a concern if the TR promoter activity is also present when delivered in a recombinant virus. Therefore, virus AAV TR-pneo was generated from TR-pneo by the University of North Carolina vector core facility by standard methods (12). AAV TR-pneo was microinjected into two regions of the brains of rats and analyzed as described by Haberman et al. (13). GFP expression was seen 2 weeks later in both regions injected: the hypothalamus and the inferior colliculus (Fig. 4). These results demonstrate that the AAV TR can generate reporter gene expression from a virus vector template as well as in plasmid form. More importantly, in vivo expression indicated that this promoter activity is not limited to tissue culture cells and may influence gene expression from rAAV vectors in the intact animal. Based on all our in vivo data established to date, we typically observed from the TR-only vectors about 2 to 5% of the level of expression seen with a wild-type CMV vector. This small amount of activity may arise as a result of AAV DNA integration 3′ of endogenous promoters. However, GFP expression was similar in all cells in all animals, which would not be expected if the expression was due to random integration. Supporting this conclusion, AAV has been demonstrated to last for long periods as episomes in muscle (5, 33). More recently, AAV TR vectors coinfected with enhancer sequences demonstrated significant expression from episomal templates in vivo (6, 29).

Transcriptional activity within terminal repeat sequences of viruses is not a novel discovery. This activity is essential for the retrovirus life cycle (32) and has been detected in other DNA viruses such as adenovirus (19, 24). Observation of transcriptional activity from the AAV TR supports previous observations seen in rabbit and monkey lung (4, 20). The exact role that this activity may play in wild-type AAV infection remains unknown. In addition, it is not clear if the activity is restricted to specific cell types or ubiquitous. We demonstrated, using plasmid and PCR linear templates, that the primary sequences necessary for activity in vitro are contained in the *D* element and the 3′ half of the *A* element. However, we did observe transcripts (only from RT-PCR) which initiated 5′ to the *A*/*D* junction, suggesting that the GC-rich secondary structure of the terminal repeat may also influence gene expression. In addition, neither of the two putative consensus initiator sequences, as described in the literature, appear to be responsible for the observed activity. Since the TR contains all of the essential functions required to generate AAV vectors (Rep/RBE binding, *trs* resolution, origin of replication, and packaging signals), we were not able to test the mini-TR expression cassettes in the context of a virus. However, our in vivo studies confirm that TR promoter activity is a very real phenomenon

that can occur in the intact animal. This observation may partly explain some of the published observations for neurotropism of AAV transduction in the brain (3). In addition, a loss of tissue-specific expression, when in the context of the AAV vector (N. Muzyczka, personal communication) may be related to TR transcriptional activity and repression. The implications of TR promoter activity on the use of AAV vectors will probably depend on the gene to be expressed, the promoter used, and the location of virus transduction in vivo. Our studies suggest that the most prominent impact of this activity will be on the regulation of gene expression. Attempts to reduce expression completely in transduced cells through the use of regulated promoters or cell-type-specific promoters may be influenced by TR promoter activity. Current vectors may have to be modified before they can be used effectively in such instances.

Although we have not mapped the AAV TR transcriptional activity to the nucleotide level, this region of the AAV TR has been extensively characterized regarding its function in replication and packaging of the viral DNA. The *A* region contains the AAV RBE and the *trs*, both of which are required for virus replication. The Rep protein binds to the terminal repeat binding element and nicks the DNA downstream at the *trs* during replication of the genome (28). The nucleotide requirements for terminal resolution have recently been mapped to a region that spans the *A*/*D* junction (2). One could try to mutate the TR in an effort to abolish the promoter activity; in fact, single-nucleotide mutations can be made without eliminating nicking of the correct site, but small decreases in nicking efficiency could have drastic effects on the ability to produce high-titer recombinant virus. In addition, spacing mutations have been generated, indicating that the correct distance from the RBE to the *trs* is crucial for AAV replication (28, 36). Results obtained with deletion mutations through the *D* sequence stress the importance of this region in packaging (35; Xiao and Samulski, unpublished). Given these data, it is unlikely that elimination of the TR promoter activity by traditional mutagenesis analysis will yield viable AAV vectors. Such mutations will probably impair viral replication and/or packaging. Since the sequences at the *A*/*D* junction are highly conserved between AAV serotypes 1 to 4 and 6, with lower conservation for type 5 (2), TR transcriptional activity may exist for all serotypes. Therefore, an underlying role for this activity may be essential for the virus life cycle (e.g., initiating wild-type virus replication or promoting templates suitable for integration). Regardless of the function in wild-type virus, this is clearly an uncontrolled activity that is part of the current AAV vectors.

It is of interest that the many strategies aimed at resolving the small packaging constraint of AAV have not considered the impact of TR promoter activity. Numerous studies are now suggesting the use of AAV "split vectors" to overcome the size limitation (6, 16, 29). The typical approach with these systems depends on one vector carrying the 5′ half of a specific gene flanked by the promoter and splice donor and another vector carrying the remaining 3′ portion of the gene flanked by the splice acceptor and a poly(A) signal. Expression of a functional gene product occurs only after recombination or reannealing through a common TR sequence from the appropriate two split gene vectors. While this and other similar strategies effectively double the packaging size of AAV, our TR data would indicate that such approaches may also increase the risk of expressing truncated gene products from TR transcriptional elements. This could lead to expression of aberrant proteins that may have transdominant negative activity, elicit an immune response, or interfere with pathways for normal secretion of the therapeutic protein (e.g., factor IX). Since the current AAV vectors are noted for safe, long-term expression with a lack of immune response, these new vectors may have to incorporate other components to eliminate the risk of unwanted expression from the TR.

Although the elimination of AAV TR promoter activity by mutation is not likely to yield viable virus, the use of insulator sequences adjacent to the TRs may provide an alternative. Insulator sequences that block the transcriptional activation of downstream promoters by enhancers have been tested in other viral vectors (31). Although adding insulator sequences to an already space-constrained vector may seem counterproductive, the significant benefit of greater control of gene expression will be necessary in some circumstances. Small insulator sequences (approximately 200 bp) have been identified that may resolve the size constraint in rAAV vectors (31). Alternatively, insertion of a poly(A) signal between the TR and the cassette of interest may blunt the unwanted expression via the TR. It appears that a number of approaches can be tested to eliminate residual transcriptional activity from the TR, and these new TR expression minus vectors should have a significant impact on the ability to utilize exogenous regulation systems and cell-type-specific promoters within an AAV vector context.

This work was supported by The Epilepsy Foundation (fellowship to R.P.H.), NINDS grant NS35633 to T.J.M., and NIH grant DK51880 to R.J.S.

We thank the UNC vector core facility for production of recombinant AAV vectors. In addition, we recognize Barrie Carter's contribution to the discovery of promoter-like activity of the AAV TRs.

REFERENCES

1. **Bohl, D., A. Salvetti, P. Moullier, and J. M. Heard.** 1998. Control of erythropoietin delivery by doxycycline in mice after intramuscular injection of adeno-associated vector. Blood **92:**1512–1517.
2. **Brister, J. R., and N. Muzyczka.** 1999. Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. J. Virol. **73:**9325–9336.
3. **Chen, H., D. M. McCarty, A. T. Bruce, and K. Suzuki.** 1998. Gene transfer and expression in oligodendrocytes under the control of myelin basic protein transcriptional control region mediated by adeno-associated virus. Gene Ther. **5:**50–58.
4. **Conrad, C. K., S. S. Allen, S. A. Afione, T. C. Reynolds, S. E. Beck, M. Fee-Maki, X. Barrazza-Ortiz, R. Adams, F. B. Askin, B. J. Carter, W. B. Guggino, and T. R. Flotte.** 1996. Safety of single-dose administration of an adeno-associated virus (AAV)-CFTR vector in the primate lung. Gene Ther. **3:**658–668.
5. **Duan, D., P. Sharma, J. Yang, Y. Yue, L. Dudus, Y. Zhang, K. J. Fisher, and J. F. Engelhardt.** 1998. Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. J. Virol. **72:**8568–8577. (Erratum, **73:**861, 1999.)
6. **Duan, D., Y. Yue, Z. Yan, and J. F. Engelhardt.** 2000. A new dual-vector approach to enhance recombinant adeno-associated virus-mediated gene expression through intermolecular cis activation. Nat. Med. **6:**595–598.
7. **Fife, K. H., K. I. Berns, and K. Murray.** 1977. Structure and nucleotide sequence of the terminal regions of adeno-associated virus DNA. Virology **78:**475–477.
8. **Flotte, T. R., S. A. Afione, R. Solow, M. L. Drumm, D. Markakis, W. B. Guggino, P. L. Zeitlin, and B. J. Carter.** 1993. Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. J. Biol. Chem. **268:**3781–3790.
9. **Flotte, T. R., R. Solow, R. A. Owens, S. Afione, P. L. Zeitlin, and B. J. Carter.** 1992. Gene expression from adeno-associated virus vectors in airway epithelial cells. Am. J. Respir. Cell Mol. Biol. **7:**349–356.
10. **Freundlieb, S., C. Schirra-Muller, and H. Bujard.** 1999. A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. J. Gene Med. **1:**4–12.
11. **Gossen, M., and H. Bujard.** 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA **89:**5547–5551.
12. **Haberman, R. P., G. Kroner-Lux, and R. J. Samulski.** 1999. Production of adeno-associated viral vectors, p. 12.9.1–12.9.16. *In* N. C. Dracopoli, J. L. Haines, B. R. Korf, D. T. Moir, C. C. Morton, C. E. Seidman, J. G. Seidman, and D. R. Smith (ed.), Current protocols in human genetics, suppl. 23. John Wiley & Sons, Inc., New York, N.Y.
13. **Haberman, R. P., T. J. McCown, and R. J. Samulski.** 1998. Inducible long-

term gene expression in brain with adeno-associated virus gene transfer. Gene Ther. **5:**1604–1611.

14. **Im, D. S., and N. Muzyczka.** 1990. The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. Cell **61:**447–457.
15. **Kistner, A., M. Gossen, F. Zimmermann, J. Jerecic, C. Ullmer, H. Lubbert, and H. Bujard.** 1996. Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. Proc. Natl. Acad. Sci. USA **93:**10933–10938.
16. **Monahan, P. E., and R. J. Samulski.** 2000. AAV vectors: is clinical success on the horizon? Gene Ther. **7:**24–30.
17. **Qing, K., X. S. Wang, D. M. Kube, S. Ponnazhagan, A. Bajpai, and A. Srivastava.** 1997. Role of tyrosine phosphorylation of a cellular protein in adeno-associated virus 2-mediated transgene expression. Proc. Natl. Acad. Sci. USA **94:**10879–10884.
18. **Rendahl, K. G., S. E. Leff, G. R. Otten, S. K. Spratt, D. Bohl, M. Van Roey, B. A. Donahue, L. K. Cohen, R. J. Mandel, O. Danos, and R. O. Snyder.** 1998. Regulation of gene expression in vivo following transduction by two separate rAAV vectors. Nat. Biotechnol. **16:**757–761.
19. **Ring, C. J., J. D. Harris, H. C. Hurst, and N. R. Lemoine.** 1996. Suicide gene expression induced in tumour cells transduced with recombinant adenoviral, retroviral and plasmid vectors containing the ERBB2 promoter. Gene Ther. **3:**1094–1103.
20. **Rubenstein, R. C., U. McVeigh, T. R. Flotte, W. B. Guggino, and P. L. Zeitlin.** 1997. CFTR gene transduction in neonatal rabbits using an adeno-associated virus (AAV) vector. Gene Ther. **4:**384–392.
21. **Ryan, J. H., S. Zolotukhin, and N. Muzyczka.** 1996. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. J. Virol. **70:**1542–1553.
22. **Samulski, R. J., M. Sally, and N. Muzyczka.** 1999. Adeno-associated viral vectors, p. 131–172. *In* T. Friedman (ed.), The development of human gene therapy, vol. 36. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Samulski, R. J., X. Zhu, X. Xiao, J. D. Brook, D. E. Housman, N. Epstein, and L. A. Hunter.** 1991. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. EMBO J. **10:**3941–3950. (Erratum, **11:**1228, 1992.)
24. **Shi, Q., Y. Wang, and R. Worton.** 1997. Modulation of the specificity and activity of a cellular promoter in an adenoviral vector. Hum Gene Ther. **8:**403–410.
25. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57:**103–113.
26. **Smale, S. T., M. C. Schmidt, A. J. Berk, and D. Baltimore.** 1990. Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proc. Natl. Acad. Sci. USA **87:**4509–4513.
27. **Smith, F. E., and T. J. McCown.** 1997. AAV vectors: general characteristics and potential use in the central nervous system, p. 79–88. *In* E. A. Chiocca and X. O. Breakefield (ed.), Gene transfer and therapy for neurological disorders. Humana Press, Totowa, N.J.
28. **Snyder, R. O., D. S. Im, T. Ni, X. Xiao, R. J. Samulski, and N. Muzyczka.** 1993. Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. J. Virol. **67:**6096–6104.
29. **Sun, L., J. Li, and X. Xiao.** 2000. Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. Nat. Med. **6:**599–602.
30. **van Holde, K., and J. Zlatanova.** 1994. Unusual DNA structures, chromatin and transcription. Bioessays **16:**59–68.
31. **Vassaux, G., H. C. Hurst, and N. R. Lemoine.** 1999. Insulation of a conditionally expressed transgene in an adenoviral vector. Gene Ther. **6:**1192–1197.
32. **Verdin, E., and C. Van Lint.** 1995. Internal transcriptional regulatory elements in HIV-1 and other retroviruses. Cell Mol. Biol. **41:**365–369.
33. **Vincent-Lacaze, N., R. O. Snyder, R. Gluzman, D. Bohl, C. Lagarde, and O. Danos.** 1999. Structure of adeno-associated virus vector DNA following transduction of the skeletal muscle. J. Virol. **73:**1949–1955.
34. **Wadkins, R. M.** 2000. Targeting DNA secondary structures. Curr. Med. Chem. **7:**1–15.
35. **Wang, X. S., K. Qing, S. Ponnazhagan, and A. Srivastava.** 1997. Adeno-associated virus type 2 DNA replication in vivo: mutation analyses of the D sequence in viral inverted terminal repeats. J. Virol. **71:**3077–3082.
36. **Xiao, X.** 1992. Characterization of adeno-associated virus (AAV) DNA replication and integration. Ph.D. thesis. University of Pittsburgh, Pittsburgh, Pa.
37. **Xiao, X., J. Li, and R. J. Samulski.** 1996. Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. J. Virol. **70:**8098–8108.
38. **Xiao, X., W. Xiao, J. Li, and R. J. Samulski.** 1997. A novel 165-base-pair terminal repeat sequence is the sole *cis* requirement for the adeno-associated virus life cycle. J. Virol. **71:**941–948.
39. **Zolotukhin, S., M. Potter, W. W. Hauswirth, J. Guy, and N. Muzyczka.** 1996. A "humanized" green fluorescent protein cDNA adapted for high-level expression in mammalian cells. J. Virol. **70:**4646–4654.