JOURNAL OF VIROLOGY, Aug. 1994, p. 4988–4997
0022-538X/94/$04.00+0

Vol. 68, No. 8

# Identification of Linear DNA Sequences That Specifically Bind the Adeno-Associated Virus Rep Protein

DOUGLAS M. McCARTY, DANIEL J. PEREIRA, IRENE ZOLOTUKHIN, XIAOHUAI ZHOU, JOHN H. RYAN, AND NICHOLAS MUZYCZKA*

*Department of Microbiology, School of Medicine, University at Stony Brook, Stony Brook, New York 11794*

Received 15 February 1994/Accepted 26 April 1994

**We have used baculovirus-expressed Rep68 that has been purified to homogeneity to reexamine the binding properties of the Rep protein. We find that Rep68 is capable of binding to a linear DNA sequence that is contained within a 25-bp sequence of the A stem of the adeno-associated virus (AAV) terminal repeat proximal to the B and C palindromes. This has been shown conclusively by demonstrating that Rep68 could specifically bind to a synthetic oligonucleotide containing the 25-bp region in the absence of the other sequences within the terminal repeat. Rep78 was also capable of binding the A stem recognition element, as demonstrated by the fact that a DNA affinity column containing the 25-bp sequence can be used to purify Rep78. The ability to recognize the linear DNA sequence within the A stem provides a mechanism by which the Rep protein can be oriented on the terminal repeat so that only the correct strand is cut at the terminal resolution site (*trs* site) during terminal resolution. In addition, computer analysis suggests that sequences similar to the A stem element are present within the three AAV promoter regions. Electrophoretic mobility shift experiments clearly demonstrate that the p5 promoter contains a Rep binding sequence. DNase protection experiments indicate that the Rep binding sequence within the p5 promoter is located between the YY1 initiator sequence and the TATA binding site. This position immediately suggests a mechanism by which the Rep protein could act as a repressor or a transactivator of p5 transcription by interacting with either YY1 or TBP. In addition, gel shift experiments suggest that the p19 promoter also contains a Rep binding site. The presence of Rep binding sites upstream of both promoters suggests that these sites may be involved in coordinate regulation of AAV transcription. In addition, we have identified a heterologous Rep binding sequence within pBR322 DNA. A comparison of the sequences within the A stem, p5, and pBR322 binding sites suggests that a repeating GAGC motif is at least part of the Rep recognition sequence. In the accompanying report (D. M. McCarty, J. H. Ryan, S. Zolutukhin, X. Zhou, and N. Muzyczka, J. Virol. 68:4998–5006, 1994), we examine the relative affinity of Rep to the A stem site and the complete terminal repeat. Finally, we also have reexamined the ability of Rep68 and Rep78 to cut at the *trs* site in substrates that do not contain the B and C palindromes or any apparent secondary structure. Although higher amounts of enzyme are necessary, *trs* endonuclease activity can be detected on such substrates with both Rep68 and Rep78. This finding suggests that the A stem recognition element and a functional *trs* site are sufficient for site specific cutting at the *trs* site and provides a mechanism by which dimer molecules, which are not expected to have hairpins, might be processed to monomer length during AAV DNA replication. It also means that the AAV terminal repeat is a tripartite origin for DNA replication. For efficient function, three elements of the terminal repeat are necessary: the *trs* site, the A stem-binding element, and the B and C palindromes.**

The *rep* gene of adeno-associated virus (AAV) codes for four overlapping polypeptides referred to by their apparent molecular sizes of 78, 68, 52, and 40 kDa (22, 34). We and others have shown that the two larger Rep proteins, Rep78 and Rep68, are the only AAV-specific gene products which are required in vivo (6, 10) and in vitro (23) to carry out AAV DNA replication. The Rep proteins of AAV contain several activities required for the terminal resolution process by which the virus maintains its genomic ends. These include binding to the AAV terminal repeat (TR) (2, 12), a site-specific, strand-specific endonuclease activity which cuts the terminal resolution site (*trs* site) (13, 33), covalent attachment at the *trs* site (31), and an ATP-dependent DNA helicase activity (13, 14). In addition, the same proteins function as positive and negative regulators of AAV gene expression (3, 17, 19, 36). The ability of Rep to carry out these multiple functions has been difficult to understand because, thus far, specific binding has been detected only to the DNA of the palindromic TR sequence, and only when it was folded into a hairpin configuration (Fig. 1a). Yet transactivation of the AAV promoters by Rep protein does not require the presence of the TRs, suggesting that Rep interacts with sequences other than the TRs (19). Furthermore, our models of terminal resolution and of AAV DNA replication suggest that Rep should also specifically recognize linear, nonhairpinned DNA (32).

The initial interaction between Rep and AAV DNA in a viral infection is probably the recognition and binding of the TR in the hairpin configuration as shown in Fig. 1. The bound Rep protein then catalyzes a single-stranded cleavage at the *trs* site, whereupon it is covalently joined to the newly generated 5′ end (31). This leaves the 3′ end available as a primer for the synthesis of DNA complementary to the A, B, and C palindromes. This asymmetric cleavage must be directed to the bottom strand of the TR so that DNA synthesis ensues toward the end of the molecule. The mechanism by which Rep

* Corresponding author. Mailing address: Department of Microbiology, University at Stony Brook, School of Medicine, Life Sciences Building 260, Stony Brook, NY 11794-5222. Phone: (516) 632-8817. Fax: (516) 632-8891.

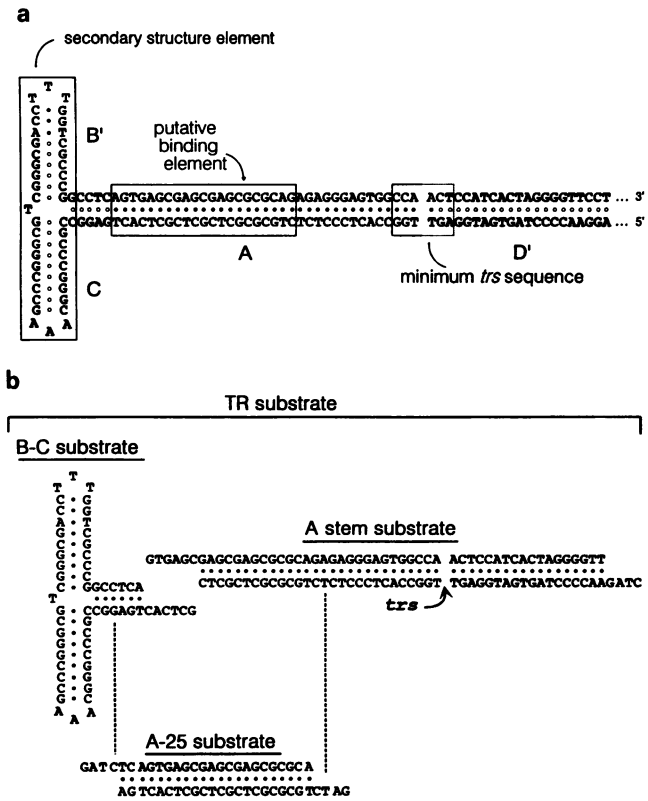


FIG. 1. Schematic representation of the AAV TR and DNA substrates for Rep binding. (a) Features of the AAV TR which are thought to interact with Rep protein. The A, B, and C palindromes and the D repeat are indicated. The boxed regions are areas which, by mutation analysis, are implicated in Rep binding and site specific endonuclease activity. Filled characters represent bases protected by Rep from DNase I digestion (12, 32). (b) Sequences of the two synthetic oligonucleotides, consisting of wild-type AAV sequences except for the *Xba*I site at the right-hand end, which were annealed to form the A stem substrate for Rep binding and *trs* activity. A third synthetic oligonucleotide, which included the B and C internal palindromes (B-C substrate), could be ligated to the A stem substrate to form a structure identical to the complete AAV TR in the hairpinned configuration. The A-25 substrate contained the 18-bp region which was common to both the A stem and the fragment of pBR322 which bound to Rep (Fig. 3). A-25 also included *Bgl*II linkers at each end which extended the sequences identical to AAV to 25 bp.

recognizes the orientation of the TR has posed a problem. Because the secondary structure of the TR is symmetric, the Rep protein must recognize one or more primary sequence elements in order to asymmetrically cut the correct strand at the *trs* site.

The process of terminal resolution inverts the orientation of the B and C palindromes relative to the *trs* site. These structures, therefore, are unlikely to provide orientation signals for Rep during terminal resolution, even though they are prominent asymmetric features of the TR. Consistent with this view, sequence substitutions in the C palindrome which maintain the secondary structure retain the ability to replicate in vivo, while those which do not allow the formation of the T-shaped structure do not replicate (4, 18, 28). This supports a role for the recognition of secondary structure and not primary sequence in the cross arms of the TR (B and C palindromes). Additional support for this interpretation comes from Rep binding studies in vitro, which demonstrate that both orientations are bound (2, 12). Finally, DNase protection assays reveal that regardless of the orientation of the TR, the small internal palindrome (B or C) further from the *trs* site is more protected. All of these considerations led us to search for a Rep-binding element within the remaining TR sequences (A and D).

One possibility for a Rep binding site was the *trs* site which links the A and D sequences. Mutations in this region demonstrated that Rep does have some capacity to recognize this sequence (32). When the *trs* site (including 3 bp to the left of the cleavage site) was moved 8 bp away from the cross arms of the T, it was still cleaved by Rep in vitro, albeit inefficiently. The altered sequence at the original *trs* position was not cleaved, suggesting specificity of the Rep protein for the primary sequence at the *trs* site. Rep also recognized and cut the *trs* site when this region of the TR was single stranded. However, further analysis of this data suggested that the *trs* site was not the only specific sequence element within the TR that interacts with Rep (32). When the *trs* region was single stranded, a significant fraction of the substrate was cut incorrectly at sites with little resemblance to sequences flanking the correct *trs* site. When only the upper ($trs^-$) strand was present, novel cleavage sites were revealed. The key point, however, was that regardless of the sequence at the *trs* site, Rep bound each of these substrates with approximately equal affinity. This observation, coupled with the fact that the D sequence was never protected led us to propose that there must be an additional Rep binding site within the A palindrome (A stem) of the TR (Fig. 1, putative binding element). Thus, our model for Rep interactions with the TR involved recognition of three elements. One was the secondary structure contributed by the B and C palindromes. Another was the *trs* site which is sequence and possibly position dependent. The last was a putative binding element within the A stem which orients Rep with respect to the *trs* site.

In addition to the role of the *rep* gene products in replication, Rep68 and Rep78 are also required for the induction of transcription from the AAV promoters (17, 36). Two lines of evidence suggest that Rep-mediated transactivation also involves binding to DNA sequences other than the TR in the hairpin configuration. First, Rep protein is capable of transactivating each of the three AAV promoters in the absence of the TRs in *cis* (19). Under these conditions, full Rep induction of the promoter at map unit 40 (p40) is dependent, in *cis*, upon sequences associated with the promoters at map units 5 and 19 (p5 and p19). Similarly, wild-type levels of p19 induction require the p5 promoter in *cis*. However, the requirement for the two upstream promoters for p40 induction is alleviated in the presence of the TRs in *cis*. We have interpreted these observations as evidence for the formation of a Rep-mediated complex between the three promoters (19). The fact that the TR, which is known to bind Rep, can substitute for the two upstream promoters suggests that direct DNA binding has a role in transactivation.

The second set of observations which implicates direct DNA binding in Rep-mediated transactivation is the activities of mutated Rep proteins. There are two separate regions of Rep which are required for TR binding in vitro (20, 25, 37). Mutations in either of these regions abolish Rep transactivation of the p19 and p40 promoters in vivo, in the absence of the TRs in *cis*. Thus, the ability to directly bind DNA correlates with the ability to transactivate.

Our previous studies of Rep binding were done with relatively small amounts of partially purified enzyme. Recently, we have purified Rep68 to homogeneity from a baculovirus expression system (23). With the availability of much larger

amounts of purified enzyme, we have reexamined the ability of Rep to bind to DNA. In this report, we demonstrate that both the Rep78 and Rep68 proteins are capable of binding to several linear duplex DNA sequences in addition to the hairpinned TR. These include the predicted site within the A stem, a related site within the p5 promoter, a sequence associated with the p19 promoter, and a sequence in heterologous DNA. In addition, we demonstrate that substrates which consist of only the A stem sequence of the TR can be nicked at the *trs* site (albeit less efficiently than the complete hairpinned TR) and can be used as an affinity matrix for the purification of Rep78. Some of these observations have also been reported recently for a mutant Rep fusion protein containing the maltose binding domain (7). In an accompanying report (21), we compare the affinity of Rep binding to the TR hairpin versus the A stem sequence and begin to characterize the binding site.

## MATERIALS AND METHODS

**Rep protein purification.** Rep68 and Rep78 proteins were extracted from recombinant baculovirus-infected SF9 insect cells as described previously (23). Rep68 protein was purified to homogeneity by chromatography on phenyl-Sepharose, single-stranded DNA (ssDNA)-cellulose, Mono S, and Mono Q as previously described (14, 23). Except for the experiments shown in Fig. 4, 5, and 8, the protein used in this study was the Mono Q fraction of baculovirus-expressed Rep68. The protein concentration was determined to be 0.22 mg/ml with the Bradford reagent (Bio-Rad), using gamma globulin as the standard. Typically, the specific activity of the Mono Q fraction for DNA replication was approximately $7.5 \times 10^5$ U/mg, using the assay of Ni et al. (23) for DNA replication. The half-life for DNA replication activity was approximately 2 weeks when the enzyme was stored at −80°C. However, binding activity, as judged by gel shift assays, was reduced only twofold and then remained constant for at least 3 months. The small loss in binding activity was apparently due to aggregation and could be reversed by mild treatment with detergent. Therefore, when indicated in the figure legends, the Mono Q fraction was treated with Tween 20 (polyoxyethylenesorbitan monolaurate) prior to its use in binding assays. This was done by mixing the enzyme with an equal volume of 1% Tween 20–0.3 M NaCl–20% glycerol–50 mM HEPES (*N*-2-hydroxyethylpiperazine-N′-2 ethanesulfonic acid)-NaOH (pH 7.9)–2.5 mM dithiothreitol (DTT)–1.5 mM $MgCl_2$ and incubating the mixture for 2 h on ice.

For comparison of *trs* endonuclease activity on A stem and TR substrates, the ssDNA-cellulose fraction of Rep68 was used. This fraction had a specific activity for DNA replication that was slightly lower than that of the Mono Q fraction but was found to be stable indefinitely at −80°C. It was also contaminated with trace amounts of a nonspecific endonuclease activity that became apparent at high enzyme concentrations. The Rep68 ssDNA fraction used had a protein concentration of 0.32 mg/ml and was greater than 95% pure (23). Rep78 *trs* endonuclease assays were also done with the ssDNA fraction for the same reason but were significantly less pure (see Fig. 5).

For the purification of baculovirus-expressed Rep78 (23), a P-11 phosphocellulose column (Whatman) was substituted for the phenyl-Sepharose chromatography step used for Rep68 in order to avoid the low-salt conditions which induce aggregation of Rep78 (data not shown). The 1 M NaCl Rep78 nuclear extract was dialyzed against buffer F (0.1 M NaCl, 50 mM Tris HCl [pH 7.5], 20% glycerol, 0.1 mM EDTA, 0.05% Nonidet P-40, 1 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride, 0.05 μg of leupeptin per ml, 0.7 μg of pepstatin per ml). The extract was then loaded onto a P-11 column which was preequilibrated with buffer F. The column was washed with buffer F and eluted with a 0.1 to 0.5 M NaCl gradient in buffer F. Rep78 protein was detected by immunoblotting as previously described (11, 13), and the Rep78-containing fractions (32 ml) were pooled and dialyzed against buffer F. The protein concentration of the pooled phosphocellulose fraction was 1.25 mg/ml with a specific activity for AAV DNA replication of $9.3 \times 10^3$ U/mg of protein, using the assay of Ni et al. (23) for DNA replication. The dialyzed phosphocellulose pool was loaded onto an ssDNA-cellulose column, which was preequilibrated with buffer F, and eluted with a 0.1 to 1.0 M NaCl gradient in buffer F. Rep-containing fractions (14 ml) were detected by immunoblotting, pooled, and dialyzed against 0.1 M NaCl in buffer R (0.25 mM Tris-HCl [pH 7.5], 20% glycerol, 2 mM $MgCl_2$, 0.1 mM EDTA, 0.05% Nonidet P-40, 1 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride, 0.5 μg of leupeptin per ml, 0.7 μg of pepstatin A per ml). The protein concentration of the dialysate was 0.275 mg/ml with a specific activity for AAV DNA replication of $9.7 \times 10^4$ U/mg. The ssDNA fraction was stable indefinitely at −80°C.

**DNA affinity chromatography.** Gel-purified A-25 oligonucleotides (Fig. 1) were annealed and ligated into concatamers. The ligated DNA (450 μg in 200 μl of water) was mixed with a slurry of cyanogen bromide-Sepharose 4B (Pharmacia) (3 g [dry weight] prewashed in 1 mM HCl) in 10 mM potassium phosphate buffer (pH 8.0). The coupling reaction mixture was incubated for 16 h at room temperature with slow rocking. The coupled DNA-Sepharose was washed with water (100 ml), and the unreacted cyanogen bromide was inactivated by incubation with 1 M ethanolamine for 3 h at room temperature. Finally, the DNA affinity column was equilibrated with 0.1 M NaCl in buffer R. For purification of Rep78, the Rep78 ssDNA-cellulose fraction described above was brought to a final concentration of 20 μg/ml with sonicated salmon sperm DNA and applied to the A-25 column in 0.1 M NaCl in buffer R at a flow rate of 1.5 to 5 ml/h. The column was washed with 5 column volumes of buffer R minus $MgCl_2$ and eluted with a linear gradient of 0.1 to 1 M NaCl in buffer R minus $MgCl_2$. The total Rep-containing volume was 1.2 ml with a protein concentration of 0.03 mg/ml. The specific activity of the pooled fraction for AAV DNA replication was $1.9 \times 10^5$ U/ml. Like the Mono Q fraction of Rep68, the affinity-purified Rep78 had a short half-life (less than 2 weeks) at −80°C.

**Electrophoretic mobility shift (EMS) assays.** The standard Rep binding reaction mixture contained, in a volume of 10 μl, 10 mM HEPES-NaOH (pH 7.9), 8 mM $MgCl_2$, 40 mM KCl, 8% glycerol, 0.2 mM DTT, 1 μg of poly(dI-dC), 0.02 to 0.1 pmol of the indicated substrate containing TR sequences, and 2 to 250 ng of Rep protein. The amounts of Rep protein and specific substrate DNA are indicated for each experiment in the figure legends. The reaction mixtures were incubated at room temperature for 20 to 30 minutes and loaded onto 4% polyacrylamide gels (29:1 acrylamide/bis-acrylamide weight ratio) containing 0.5× TBE buffer (45 mM Tris, 89 mM boric acid, 1 mM EDTA). The gels were electrophoresed at room temperature in 0.5× TBE buffer for 2 h at 9 V/cm.

The binding reactions with the p5 and p19 promoter substrates were performed in a volume of 20 μl containing 20 mM HEPES (pH 7.9), 5 mM $MgCl_2$, 50 mM NaCl, and 1 μg of poly(dI-dC). Probes for the p5 and p19 promoter binding assays were generated by PCR to amplify sequences from either plasmid pIM45 (19) or a plasmid with a deletion between the p5 and p19 promoters, p5-19. In addition, the

*Bgl*II-*Bam*HI fragment from p5-19 was 3′ labeled and used as a probe in some experiments. Plasmid p5-19 was made from pIM45 by oligonucleotide-directed mutagenesis (15), using a primer which replaced sequences from AAV nucleotides 320 to 671 with an *Xho*I restriction site (oligonucleotide sequence, 5′-CGCAGCCGCCCTCGAGTCAGAGAATT-3′). The PCR mixtures contained, in a volume of 75 μl, 50 mM KCl, 10 mM Tris-HCl (pH 8.8), 3 mM $MgCl_2$, 0.1% Triton X-100, 1.5 mM each dATP, dCTP, dGTP, and dTTP, 0.014 pM plasmid DNA, 100 pM $^{32}$P-5′-end-labeled oligonucleotide, 100 pM unlabeled oligonucleotide, and 2 U of *Taq* DNA polymerase (Promega Biotec). The reactions were cycled 27 times through 45 s at 94°C, 45 s at 59°C, and 30 s at 72°C. The probes were subsequently gel purified. The primers used for PCR amplification of probes were 20 bases long and are identified in Results by the position of their 5′ nucleotide.

***trs* assays.** Assays for *trs* endonuclease activity were performed under the same reaction conditions as the EMS assays described above for the TR substrates except that ATP was added to a final concentration of 0.5 mM and the reaction mixtures were incubated for 2 h at 37°C. The reaction mixtures contained a substrate concentration of 1 pmol/ml in a final volume of 10 μl and either 17 to 275 ng of the ssDNA fraction of Rep78 or 20 to 320 ng of the ssDNA fraction of Rep68. These *trs* endonuclease assay conditions were different in several respects from those used previously (13), the major differences being the use of a fourfold-higher substrate concentration, the absence of bovine serum albumin, and a higher Rep concentration. The reactions were stopped by the addition of 10 μl of 0.6 M sodium acetate–20 mM EDTA–1% sodium dodecyl sulfate (SDS) and precipitated with 2.5 volumes of ethanol. The precipitates were dissolved in formamide loading buffer, incubated for 10 min at 68°C, and electrophoresed on an 8% acrylamide–7 M urea sequencing gel.

**DNase I footprinting assays.** The footprinting assays using plasmid DNA were based on the methods used by Gralla (8). The Rep binding reaction mixtures were exactly as detailed above and contained 50 or 100 μg of Rep Mono Q fraction per ml. Following the binding step, the reaction mixtures were incubated for 3 min at 37°C. Each reaction mixture was then mixed with 5 μl (prewarmed to 37°C) of 20 mM HEPES-NaOH (pH 7.9)–8 mM $MgCl_2$–4 mM $CaCl_2$–2 U of DNase I (Worthington) per ml and incubated for 1 min at 37°C. The DNase digestion was stopped by the addition of 50 μl of a solution containing 0.6 M sodium acetate, 20 mM EDTA, and 1% SDS, and the final volume was brought to 100 μl with $H_2O$. The reaction was then extracted with phenol and chloroform (1:1) and precipitated with 2.5 volumes of ethanol. The precipitate was dissolved in 10 μl of 50 mM Tris-HCl (pH 7.2)–10 mM $MgCl_2$ to which was added 1 μl of 250 pmol of $^{32}$P-end-labeled primer per ml. The DNA was heated for 4 min at 95°C and then annealed at 55°C for 15 min. After a brief centrifugation at room temperature, primer extension was carried out by the addition of 5 μl of 1 mM (each) dATP, dCTP, dGTP, and dTTP, 2.0 mM DTT, 100 μg of bovine serum albumin per ml, and 500 U Sequenase version 2.0 T7 DNA polymerase (United States Biochemical) per ml. Reaction mixtures were incubated for 30 min at 37°C, precipitated with ethanol, and electrophoresed on 6% acrylamide–7 M urea sequencing gels.

DNase I footprinting of the TR in the hairpin configuration was done in a reaction identical to the binding and DNase I digestion assays described above except that the reaction mixture contained, per milliliter, approximately 0.5 pmol of either 5′- or 3′-end-labeled *Xba*I fragment of NE DNA (28) and 12.0 or 24.0 μg of Rep68 Mono Q fraction. The DNase I digestion products were analyzed on an 8% acrylamide–7 M urea sequencing gel.

For DNase I footprinting of the p5 promoter, Rep was bound to the labeled DNA in a 50-μl reaction mixture containing 25 mM HEPES (pH 7.9), 5 mM $MgCl_2$, 50 mM NaCl, 2% glycerol, and 20 μg of poly(dI-dC) per ml. After 20 min at room temperature, 50 μl of a solution containing 10 mM $MgCl_2$ and 5 mM $CaCl_2$ was added to the reaction mixture, and it was placed on ice for 5 min. One microliter of DNase I (100 U/ml) was added, and incubation was continued on ice for 2 min. The reaction was stopped by the addition of 90 μl of 20 mM EDTA–1% SDS–0.2 M NaCl–1.1 mg of carrier tRNA per ml and then extracted with phenol and chloroform (1:1) and precipitated with ethanol. The DNase digestion products were analyzed on a sequencing gel.

## RESULTS

**Discrete Rep binding sites within the TR A palindrome.** As mentioned earlier, previous studies have demonstrated the binding of Rep protein only to the AAV TR in the hairpin configuration. To detect Rep binding to specific sequences in the absence of secondary structure, we assayed for Rep-DNA interactions on a plasmid, pSM620 (27), which contained the entire AAV genome cloned into pBR322. The TR sequences, in the context of this plasmid, were in linear duplex form rather than a hairpin or cruciform configuration (38). Rep68 binding was detected by the DNase protection assay of Gralla (8). The putative Rep68-DNA complex was treated with DNase I, and the protected region was visualized by primer extension. The primer that was used was oriented leftward from the unique region of the viral genome just upstream from the p5 promoter (AAV nucleotides 237 to 218).

The Rep68 Mono Q fraction strongly protected three small regions within the A palindrome of the AAV TR (Fig. 2). The first included AAV nucleotides 93 to 96, and the second included nucleotides 101 to 104. Each of these protected regions comprised the tetrameric sequence GAGC. An additional copy of GAGC separated the two regions. The third area of protection extended from nucleotides 119 to 122 (TGGC), which were 2 bp to the left of the *trs* site (nucleotide 124). We referred to these as regions 1, 2, and 3, from left to right (Fig. 2). We also observed a more extended, though weaker, protection over most of the A palindrome. There was no evidence for protection of the B or C palindrome.

This pattern was qualitatively different from previously observed footprints on the AAV TR (2, 12). In the earlier studies, the TR in the hairpin configuration had been used as the substrate, and this consistently resulted in complete protection of most of the A palindrome and parts of the B or C palindrome, depending on the orientation of the TR (Fig. 1). Our assay also differed from earlier binding studies in the source of the Rep protein, which had been isolated from AAV-infected HeLa cells in previous experiments. To determine whether this footprint was a characteristic of the insect cell-expressed Rep, we used the Rep68 Mono Q fraction in DNase protection assays with the TR in the hairpin configuration. The DNA substrate was the same as had been used in previous studies, the NE *Xba*I fragment (12), and was labeled at either the 5′ or 3′ end. The pattern of DNase protection with this substrate in the presence of the baculovirus-expressed Rep68 was the same as that seen previously with Rep68 obtained from HeLa cells (data not shown). This finding suggested that the highly localized footprint observed in the plasmid protection assay was related to the conformation of the DNA substrate and not the source of the enzyme. These

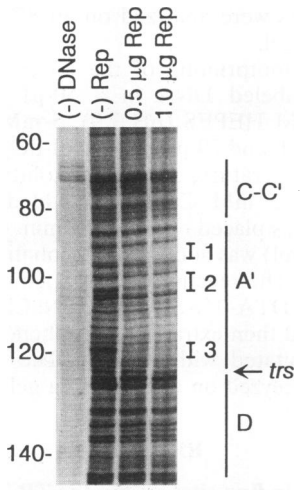


FIG. 2. DNase I footprint analysis of the AAV TR region in supercoiled plasmid DNA. Plasmid pSM620 (27) (0.5 μg), containing the entire AAV genome, was bound to the indicated amount of purified Rep68 protein (Mono Q fraction) and digested with DNase I. Regions of protection were visualized by primer extension with the AAV coding strand as the template and a primer complementary to nucleotides 237 to 218. AAV nucleotide positions are indicated on the left, and structural features of the TR are shown on the right. Protected regions are highlighted by bars on the right and are labeled 1, 2, and 3. The four reactions included plasmid with Rep in the absence of DNase treatment [(−) DNase], treatment of naked plasmid [(−) Rep], and two concentrations of Rep protein followed by DNase treatment (0.5 and 1.0 μg Rep).

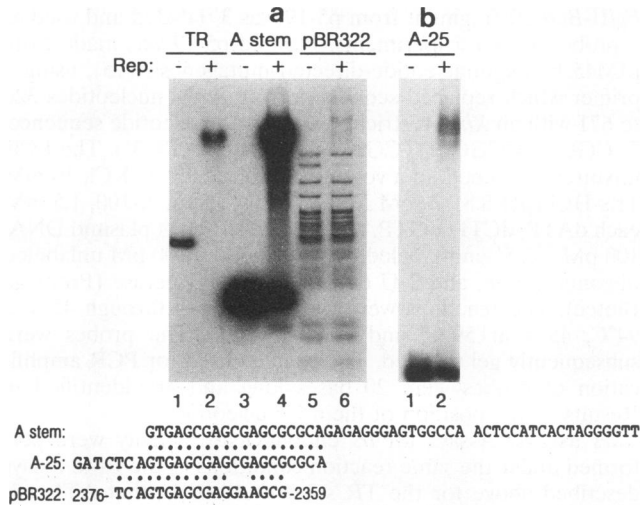


FIG. 3. EMS assays for Rep binding. Four different DNAs were used to test Rep substrate binding as described in Materials and Methods. Each is shown in the presence (+) or absence (−) of Rep protein. The TR lanes contained 0.014 pmol of 3′-labeled terminal *Xba*I fragment from NE DNA (12). The pBR322 lanes contained 0.4 nmol of an *Hpa*II digest of pBR322. The EMS assay for the A-25 substrate is from a separate experiment which was done with 1.5 pmol of the substrate. All experiments were done with 110 ng of Rep68. An alignment of the top strands of the A stem and A-25 substrates with the relevant region of pBR322 is shown below.

observations also suggested that Rep protein could recognize and bind to linear duplex DNA sequence and that the recognition element for binding was within the A palindrome of the TR.

**Rep binds specifically to a subsequence of the A palindrome in the AAV TR.** To test the idea that Rep could bind to the A palindrome in the absence of secondary structure, we did EMS assays with a synthetic double-stranded substrate that contained most of the A sequence (including the *trs* site) and most of the D sequence (Fig. 1, A stem substrate). Because it did not contain the B and C sequences, the A stem substrate could not adopt a hairpinned or cruciform-like structure. Nevertheless, the A stem substrate was clearly capable of binding to Rep68 (Fig. 3a, lanes 3 and 4) under the same conditions as the complete hairpinned TR (Fig. 3a, lanes 1 and 2). In this experiment, the TR and the A stem were not normalized for either concentration or specific activity, and so the intensity of the A stem-protein complex reflects the use of higher concentrations of this substrate and possibly its higher specific activity. In addition, an *Hpa*II digest of pBR322 was included as a control in this experiment to confirm that the Rep binding reaction was sequence specific (Fig. 3a, lanes 5 and 6). Indeed, most of the pBR322 *Hpa*II fragments were not bound by Rep68. However, one of the fragments, *Hpa*II B, clearly shifted in the presence of Rep protein. A comparison of the sequences of AAV and pBR322 revealed a 16-of-18 base match between the AAV A palindrome and the *Hpa*II B fragment of pBR322. The sequence alignment is shown at the bottom of Fig. 3 and in Table 1 (AAV nucleotides 88 to 105, pBR322 nucleotides 2376 to 2359). This fortuitous similarity immediately suggested that this subsequence was sufficient for Rep binding. To test this possibility, we synthesized oligonucleotides which contained 25 bp of the AAV sequence centered on the region which was common to pBR322 and the A stem (Fig. 1, A-25 substrate). The ends of the duplex oligonucleotide formed *Bgl*II linkers which were filled in for the EMS assays. The experiment in Fig. 3b demonstrated that this oligonucleotide was also efficiently bound by the baculovirus Rep68 protein. These 25 bp apparently were sufficient for stable Rep binding. We note that the 25-bp sequence is virtually identical to the putative binding element predicted by Snyder et al. on the basis of indirect data (Fig. 1) (32). In addition, the A-25 substrate contained two of the regions identified in the DNase protection experiment (Fig. 2, regions 1 and 2).

**Affinity purification of Rep78.** We reasoned that if Rep could bind specifically to the A-25 oligonucleotide, it should be possible to use the oligonucleotide to construct an affinity column for the purification of the Rep protein. To test this idea, the oligonucleotide containing the 25-bp Rep binding sequence was ligated into concatemers and coupled to a cyanogen bromide-activated Sepharose matrix. The starting material for purification was a 1.0 M nuclear extract from insect cells infected with a recombinant baculovirus expressing wild-type Rep78 protein. The extract was first fractionated by phosphocellulose and ssDNA-cellulose chromatography as described in Materials and Methods. The ssDNA-cellulose fraction was then applied to the DNA affinity matrix in the presence of 0.2 M NaCl and 2.0 mM $MgCl_2$ and eluted with a gradient of 0.2 to 0.6 M NaCl (Fig. 4). The affinity column fractions were assayed by immunoblotting, and the Rep-containing fractions were also assayed for *trs* endonuclease activity (Fig. 4a and b, respectively).

A significant fraction of the Rep78 bound to the A-25 column (Fig. 4a). When the bound fraction was eluted from the column, it contained *trs* endonuclease activity (Fig. 4b). Furthermore, a comparison of the affinity-purified Rep78 with the ssDNA fraction by SDS-acrylamide gel electrophoresis indicated that a significant purification had been achieved (Fig.

TABLE 1. Known and putative Rep68 binding sites

| Site context | Sequence[a] | AAV nucleotides | Rep binding |
|---|---|---|---|
| A stem (A-25) | CTCAGTGAGCGAGCGAGCGCGCAGA | 87–111 | + |
| pBR322 | gTCAGTGAGCGAGgaAGCGgaagag | 2377–2353[b] | + |
| p5 | aagcccGAGtGAGCacGCagGgtct | 260–284 | + |
| p19 | tcCAGTGgGCGtGgactaatatgGA | 814–838 | +[c] |
| | cTCAcgGAGCGtaaacGgttGgtGg | 864–888 | +[c] |
| p40 | aTaAGTGAGCccaaacGgGtGCgcg | 1824–1848 | ND[d] |
| | acggGTGcGCGAGtcAGttgcgcag | 1838–1862 | ND |
| Unknown | agaAGgGAGaGAGCtActtCcacat | 568–592 | ND |
| | aggccgGAGCGAGtGAcattcggGA | 3593–3617 | ND |
| HIV long terminal repeat | CctAGTtAGCcAGaGAGCtCcCAGg | 29–53[e] | ND |

[a] Nucleotides indicated in uppercase are an exact match with the A stem 25-bp sequence.
[b] From the pBR322 sequence.
[c] A DNA fragment containing both of these sites is bound by Rep68; it is not known which site is bound.
[d] ND, not determined.
[e] From the HIV-1 sequence; adapted from references 24 and 26.

5). In addition to Rep78, the affinity-purified fraction contained only two other prominent bands, of approximately 215 and 35 kDa. The identities of these copurifying proteins were not determined. The specific activity of the affinity-purified Rep78 fraction was approximately $1.9 \times 10^5$ U/mg (data not shown), as determined by the replication assay of Ni et al. (23). This was approximately 20% of the highest specific activity that we have achieved thus far for Rep68 and represented a twofold purification over the ssDNA fraction.

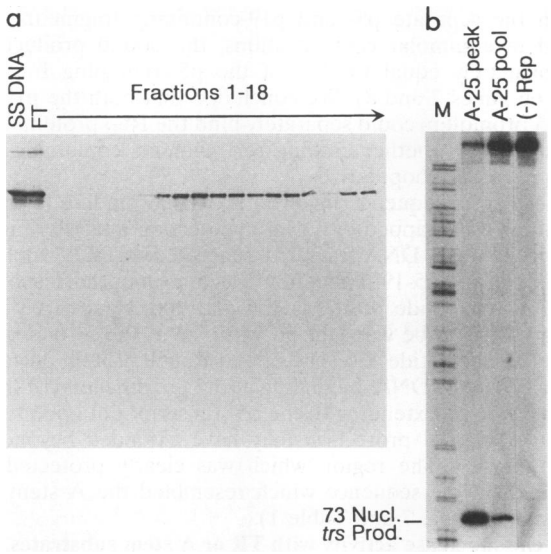


FIG. 4. Purification of Rep78 by an A-25 DNA affinity column. (a) Aliquots of fractions eluted from an A-25 DNA affinity column, as well as the starting ssDNA-cellulose fraction (SS DNA) and the flowthrough fraction (FT), were immunoblotted with an anti-52/40 Rep monoclonal antibody (11). The column was eluted with a 0.1 to 1.0 M NaCl gradient as described in Materials and Methods. (b) *trs* endonuclease assay of the affinity-purified Rep78. The substrate is the 266-nucleotide terminal *Xba*I fragment of NE DNA; the expected Rep cleavage product is a 73-nucleotide (Nucl.) fragment as indicated. The molecular weight marker (M) is an *Hpa*II digest of pBR322. Assays include a negative control [(−) Rep], untreated fraction 7 (A-25 peak), and pooled and concentrated Rep-containing fractions (A-25 pool).

The large amount of Rep in the flowthrough fractions (Fig. 4a, FT lane) was typical. Incomplete binding was apparently not due to simple overloading of the affinity column (data not shown). In the accompanying report (21), we demonstrate in more detail that we were unable to find substrate and enzyme concentrations at which all of the Rep protein is bound to the A stem sequence.

The use of the A-25 affinity column for Rep78 purification demonstrates that Rep78, as well as Rep68, can bind to the A-25 sequence. The fact that a significant purification of Rep78 was achieved by affinity chromatography, and the fact that Rep did not bind to nonspecific double-stranded DNA columns (data not shown), confirms that binding to the A stem

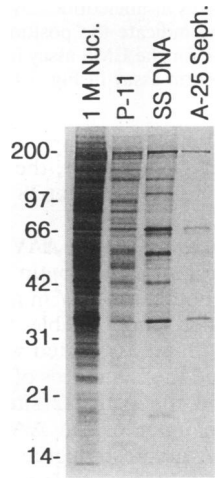


FIG. 5. Purification of Rep78. Nuclei from SF9 cells infected with recombinant baculovirus expressing Rep78 were extracted with 1 M NaCl, and the extracts were fractionated by conventional and affinity chromatography. Aliquots from each step of the Rep78 purification were electrophoresed on an SDS–7 to 14% gradient acrylamide gel and visualized by silver stain. The crude extract (1 M Nucl.), phosphocellulose fraction (P-11), ssDNA-cellulose fraction (SS DNA), and A-25 DNA affinity fraction (A-25 Seph.) are shown. The molecular masses (in kilodaltons) and positions of molecular weight markers are indicated on the left.

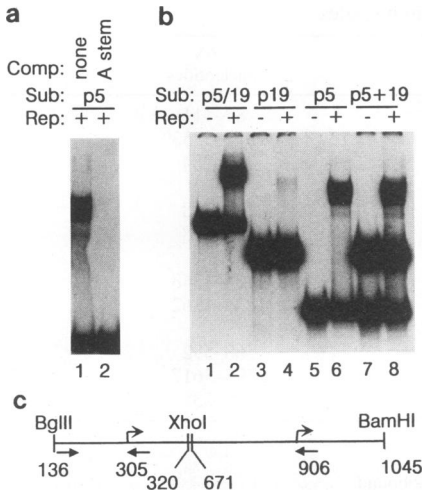


FIG. 6. Rep68 protein binds to the AAV p5 and p19 promoters. (a) A 170-bp 5′-end-labeled DNA fragment (AAV nucleotides 136 to 305) containing the p5 promoter was synthesized by PCR using the primers indicated in panel c. Approximately 0.013 pmol of the fragment was incubated with Tween 20-treated Rep68 and electrophoresed on a 4% acrylamide gel. See Materials and Methods for additional details. Reactions with (+) and without (−) Rep protein are indicated. The Rep-p5 complex is visible as the upper band in lane 1. Competition (Comp) from a 300-fold excess of the A-25 oligonucleotide substrate (Sub) (Fig. 1) eliminated binding to the p5 fragment in lane 2. (b) The 560-bp *Bgl*II-*Bam*HI fragment illustrated in panel c was isolated from plasmid p5-19, labeled at its 3′ end, and used as a probe for Tween 20-treated Rep68 binding as described in Materials and Methods and Results (lanes 1 and 2). In addition, portions of the *Bgl*II-*Bam*HI fragment containing only the p5 or p19 promoter were isolated by *Xho*I digestion and used as probes, either separately (lanes 3 and 4, p5 promoter; lanes 5 and 6, p19 promoter) or as an equimolar mixture (lanes 7 and 8). (c) The *Bgl*II-*Bam*HI fragment in plasmid p5-19 containing AAV nucleotides 136 to 1045. XhoI indicates the position of the deletion between nucleotides 320 and 671 and the *Xho*I linker insertion. The bent arrows indicate the positions of the p5 (left) and p19 (right) mRNA start sites at nucleotides 287 and 873, respectively. The arrows below the line indicate the positions of primers used for generating probes by PCR for the EMS assay in panel a (136 and 305) and for the DNase protection assay in Fig. 7 (136 and 906).

element is sequence specific. Finally, the affinity column may be useful for the purification of Rep78 and Rep-associated cellular proteins.

**Rep binds to sequences within the AAV p5 and p19 promoters.** As mentioned earlier, Rep protein had been shown to regulate the three promoters of AAV in a complex manner (3, 17, 19, 36). It was therefore possible that additional Rep binding sites could be found associated with one or more the promoters. We first used the sequences of the Rep binding site within the A stem and the pBR322 site as a consensus to compare with the remainder of the AAV genome. The sequences displaying the greatest degree of similarity to the consensus were located in the regions of the three AAV promoters (Table 1). The largest such related sequence was located between the TATA site and the initiation site of the p5 promoter (AAV nucleotides 266 to 279).

To determine whether the A stem-related sequence in the p5 promoter was a functional Rep binding site, we tested a p5-containing DNA fragment, consisting of sequences from AAV nucleotides 136 to 305, for binding to the Rep68 Mono Q fraction. We generated the probe for this experiment by PCR amplification from plasmid pIM45 (19) (Fig. 6c). The binding of Rep68 to this fragment was demonstrated in the assay shown in Fig. 6a, lane 1. The protein-DNA complex was inhibited by an excess of unlabeled A stem substrate (Fig. 6a, lane 2), suggesting that the p5 and A stem sites were functionally related. This was the first observation of Rep binding in the context of an AAV promoter.

We next wanted to test for Rep binding to the p19 promoter. We had proposed previously (19) that Rep induction of the p19 and p40 promoters involved the formation of a complex between the p5 promoter and the other promoters in *cis*. We therefore tested the ability of the p19 promoter to bind Rep, either by itself or linked to the p5 promoter. In this experiment, restriction fragments of the deletion plasmid p5-19 were used as probes (Fig. 6c). Plasmid p5-19 (see Materials and Methods) contained a deletion of AAV nucleotides 321 to 670, thus bringing the p5 and p19 promoters closer together. In addition, an *Xho*I linker was inserted at the site of the deletion so that binding could be assayed with the two promoters in *cis* or on separate fragments. The 560-bp *Bgl*II-*Bam*HI fragment isolated from the plasmid was 3′ labeled and contained both promoters (Fig. 6c). The p5 and p19 mRNA start sites (at nucleotides 287 and 873, respectively) were 242 bp apart in this fragment. Digestion of the *Bgl*II-*Bam*HI fragment with *Xho*I produced a 195-bp p5-containing probe and a 375-bp p19-containing probe.

The *Bgl*II-*Bam*HI fragment (Fig. 6b, lanes 1 and 2) bound Rep as expected since it contained the p5 binding site. The level of binding to this probe, in which the two promoters were linked, was not significantly greater than the binding to the probe containing only the p5 promoter (Fig. 6b; compare lanes 2 and 6). There was, however, a low level of binding to the p19 probe alone (Fig. 6b, lanes 3 and 4). Binding to the p19 fragment was specific because it was inhibited by the A stem fragment but not a heterologous fragment (data not shown). When the separate p5- and p19-containing fragments were mixed in equimolar concentrations, the bound product was approximately equal to that of the p5-containing fragment (Fig. 6b, lanes 7 and 8). We concluded that both the p19 and the p5 promoters could separately bind the Rep protein, but it was not clear whether binding to fragments containing both promoters was cooperative.

The precise sequence involved in Rep68 binding to the p5 promoter was mapped by DNase protection analysis (Fig. 7). The probe was a DNA fragment generated by PCR amplification of plasmid p5-19. The left and right primers corresponded to AAV nucleotide positions 136 and 906, respectively (Fig. 6c), and the probe was labeled with $^{32}P$ at the 5′ end of the primer at nucleotide 136. The bound Rep68 protein protected at least 26 bp of DNA, beginning at the proximal end of the p5 TATA site and extending to the p5 transcript initiation site. A region of partial protection may have extended beyond the initiation site. The region which was clearly protected was centered on the sequence which resembled the A stem Rep binding site (Fig. 7 and Table 1).

***trs* endonuclease activity with TR or A stem substrates.** The Rep68 and Rep78 proteins had been shown to catalyze the single-stranded endonucleolytic cleavage of the *trs* site when a TR in the hairpin configuration was used as substrate. The A stem oligonucleotide contained both the Rep binding sequence and the *trs* site. To determine whether this was sufficient for binding and *trs* cleavage (or whether sequences within the B and C palindromes were required), we compared the *trs* endonuclease activities of both Rep78 and Rep68 on the TR and A stem substrates. For this experiment, we used the ssDNA-cellulose fractions of baculovirus-expressed Rep78 and Rep68. As mentioned earlier (Materials and Methods), these

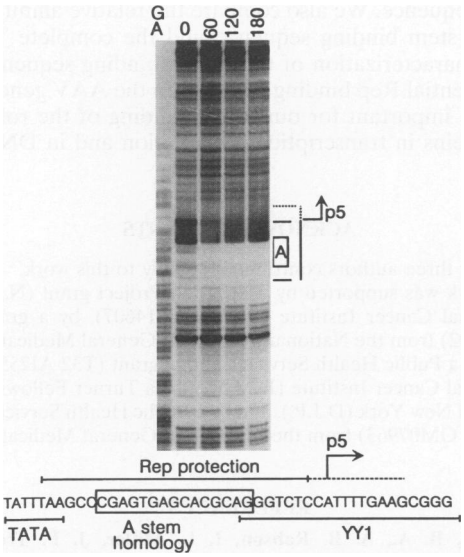


FIG. 7. DNase I footprint of Rep bound to p5 promoter. A DNA fragment from plasmid p5-19 was amplified by PCR using primers at nucleotide positions 136 and 907 (see Fig. 6c and Materials and Methods). The amplified fragment was labeled with $^{32}P$ at the 5′ end of the nucleotide 136 position. The probe was incubated with Tween 20-treated Rep68 protein and then digested with DNase I. The digestion products were electrophoresed on a 6% acrylamide sequencing gel. The protected region is indicated by a solid line at the right. The broken line, beginning at the indicated p5 mRNA start site, represents a region of partial protection. The box labeled A spans the sequence related to the A stem Rep binding site. The DNA sequence of the p5 region is shown below. The protected region, the A stem homology, and the p5 promoter are indicated as above. The positions of the TATA site and a YY1 transcription factor binding site (31, 30, 34) are also indicated.

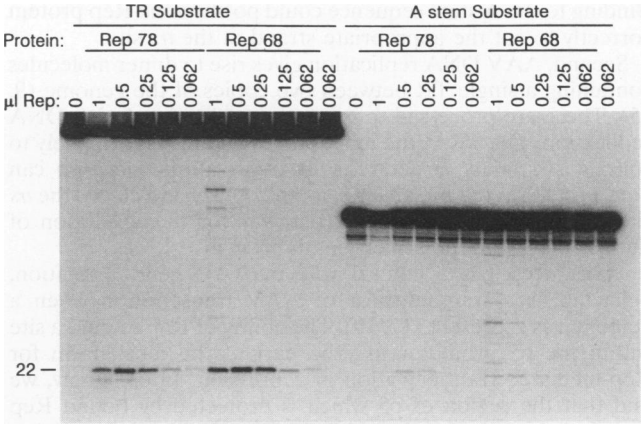


FIG. 8. Comparison of *trs* endonuclease activity with TR and A stem substrates. Each substrate (0.02 pmol) was incubated in volume of 10 μl in the presence of 0.5 mM ATP and twofold dilutions of the ssDNA-cellulose fractions of Rep78 (17 to 275 ng) or Rep68 (20 to 320 ng). See Materials and Methods for additional details. Specific cutting of either substrate at the *trs* site generates the indicated 22-nucleotide product.

fractions were relatively stable for DNA replication and *trs* activity upon storage, whereas more highly purified preparations rapidly lost catalytic activity while retaining binding activity.

Each substrate was incubated with increasing amounts of either Rep78 or Rep68 (Fig. 8). As expected, both Rep68 and Rep78 were capable of cutting the *trs* site in the complete hairpinned TR substrate to produce the expected 22-base product. We also detected *trs* endonuclease activity on the A stem substrate with both Rep78 and Rep68. We estimated that the level of *trs* endonuclease activity with the A stem substrate was approximately 5- to 10-fold lower than that of the complete TR. Nevertheless, it was clear that a substrate that was missing the B and C palindromes but retained the Rep binding element and the *trs* site could be correctly cut by both Rep68 and Rep78.

Additional cleavage sites were detected with both enzymes at the highest enzyme concentrations. It was not clear whether this was due to contaminating nuclease activity that was present in these fractions or incorrect cutting by Rep protein. In addition, the highest concentrations of Rep appeared to be inhibitory. Finally, in the case of Rep68, we estimated that the enzyme-to-substrate ratio was 16:1 to 260:1 over the range of enzyme concentrations that was tested. This was significantly higher than the approximately 7:1 ratio that we have used previously to detect *trs* endonuclease activity (13).

## DISCUSSION

We have identified a linear Rep binding sequence which is within the A stem region of the AAV terminal repeat. Although we had predicted the existence of this binding element, our previous analyses had detected Rep binding only to the complete TR in the hairpin configuration (12, 32). These studies now confirm the location of the predicted binding element (32). Our present observation of Rep binding to the TR in the context of a supercoiled plasmid indicates that Rep can recognize the A stem portion of the TR in a linear duplex conformation without making contacts with the B and C palindromes. This is conclusively demonstrated through binding experiments using the A stem in the absence of the B and C palindromes as the substrate. Further, the isolated A stem is a substrate for cleavage at the *trs* site by either Rep68 or Rep78 protein, although at a lower frequency than the complete TR. We note that similar observations have been made recently for a mutant Rep fusion protein containing the maltose binding domain (7). These observations, along with our previous work (32), clearly demonstrate that the AAV TR is a tripartite origin for DNA replication which requires three elements for efficient function: the *trs* site, the A stem-binding element, and the B and C palindromes. We have also found that Rep protein binds to a site within the AAV p5 promoter which resembles in sequence the A stem binding site (Table 1). An additional sequence capable of binding Rep was found in the p19 promoter, but this sequence, because of its lower affinity, has not yet been precisely mapped. Finally, a sequence capable of binding Rep was found in heterologous pBR322 DNA.

The identification of this linear Rep binding sequence solves several of the problems that we have encountered in understanding the multiple roles of Rep protein. First, Rep must recognize the AAV TR in the hairpin configuration and cleave asymmetrically on the appropriate strand during the process of terminal resolution. Because of the symmetry of the TR, it was not clear how Rep was oriented correctly to cut only one strand at the *trs* site. The A stem recognition sequence, which is intrinsically asymmetric, potentially solves this problem, and this was our primary reason for predicting its existence (32).

Binding to the A stem sequence could position the Rep protein correctly to cut the appropriate strand at the *trs* site.

Second, AAV DNA replication gives rise to dimer molecules containing a single TR between two copies of the genome (9, 35). These are processed to monomer molecules during DNA replication. The TR at the junction of the dimer is not likely to contain secondary structure. Our observation that Rep can bind to a TR in the linear state and recognize and cleave the *trs* site (albeit poorly) provides a pathway for the resolution of these dimers into monomer-length progeny.

Third, Rep has a central role in AAV gene regulation, behaving as a transactivator of AAV transcription when a helper virus is present (17, 19). The ability of Rep to bind a site within the p5 promoter may be part of the mechanism for Rep-mediated transactivation or repression. Interestingly, we find that the region of p5 which is protected by bound Rep protein is between a known binding site for the transcription factor YY1 and the TATA site (5, 29, 30, 34). Thus, the position of the Rep binding site suggests that Rep might make contacts with YY1 or the TATA-binding protein. YY1 has been demonstrated to mediate positive and negative regulation of the AAV p5 promoter in the presence and absence, respectively, of adenovirus E1A (5, 29, 30). The presence of the Rep binding site within the p5 promoter suggests that Rep, YY1, and the E1A protein might cooperate in the regulation of the p5 promoter.

Fourth, we showed previously that Rep transactivation of the AAV p19 and p40 promoters, in the absence of the TRs, depends on the presence of both the p5 and p19 promoters in *cis*. We suggested that Rep binds to all three promoters and forms a Rep-mediated complex between the three promoters by DNA looping. The fact that we have now demonstrated Rep binding to both the p5 and p19 promoters supports this idea. Furthermore, computer analysis of the p40 promoter revealed that it also contains two sequences that may bind Rep, although we have not yet tested this possibility (Table 1). If, indeed, Rep binds to all three AAV promoters, this would provide a simple mechanism by which Rep could coordinately transactivate the three promoters.

Finally, a number of laboratories have demonstrated that Rep negatively regulates the expression of heterologous gene expression (1, 16, 24, 26). An example of this is the inhibition of human immunodeficiency virus (HIV) infection by AAV (1, 24, 26). This effect may be due to interference with the HIV Tat protein interaction with the HIV long terminal repeat. Sczakiel and colleagues (24, 26) have identified a similarity between the HIV long terminal repeat and the A stem region that binds Rep (24, 26) (Table 1). The ability of Rep to bind to this sequence, therefore, would provide a direct mechanism for the inhibition of HIV gene expression.

The fortuitous observation that a fragment of pBR322 would bind to Rep protein suggested that an 18-bp sequence which is similar (89% identical) to that of the A stem is sufficient for Rep binding. Comparison of the 25-bp A stem binding sequence, the probable pBR322 binding sequence, and the p5 binding sequence (Table 1) suggests that one or more copies of a common GAGC sequence (or a related tetramer) may be essential for binding. The DNase I protection pattern of Rep68 protein bound to the TR in a supercoiled plasmid also suggests that the tandemly repeated GAGC sequence is at least part of the recognition element for binding. Two of the four GAGC (or GCGC) tetramers were protected from DNase I digestion in this assay, which might suggest that the GAGC motifs are multiple independent Rep sites. In the accompanying report (21), we demonstrate by mutational analysis that this is probably not the case and provide additional information about the binding sequence. We also compare the relative affinity of Rep to the A stem binding sequence and the complete TR. The further characterization of these Rep binding sequences, and other potential Rep binding sites within the AAV genome, will clearly be important for our understanding of the role of the Rep proteins in transcriptional regulation and in DNA replication.

## ACKNOWLEDGMENTS

The first three authors contributed equally to this work.

This work was supported by a Program Project grant (N.M.) from the National Cancer Institute (PO1 CA2814607), by a grant (RO1 GM3572302) from the National Institute of General Medical Sciences (N.M.), by a Public Health Service training grant (T32 AI25530) from the National Cancer Institute (D.M.M.), by a Turner Fellowship from the State of New York (D.J.P.), and by a Public Health Service training grant (T32 GM07963) from the Institute of General Medical Sciences (J.H.R.).

## REFERENCES

1. **Antoni, B. A., A. B. Rabson, I. L. Miller, J. P. Trempe, N. Chejanovsky, and B. J. Carter.** 1991. Adeno-associated virus Rep protein inhibits human immunodeficiency virus type 1 production in human cells. J. Virol. **65**:396–404.
2. **Ashktorab, H., and A. Srivastava.** 1989. Identification of nuclear proteins that specifically interact with adeno-associated virus type 2 inverted terminal repeat hairpin DNA. J. Virol. **63**:3034–3039.
3. **Beaton, A., P. Palumbo, and K. I. Berns.** 1989. Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in *trans* by the Rep protein. J. Virol. **63**:4450–4454.
4. **Bohenzky, R. A., R. B. LeFebvre, and K. I. Berns.** 1988. Sequence and symmetry requirements within the internal palindromic sequences of the adeno-associated virus terminal repeat. Virology **166**:316–327.
5. **Chang, L. S., Y. Shi, and T. Shenk.** 1989. Adeno-associated virus p5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. J. Virol. **63**:3479–3488.
6. **Chejanovsky, N., and B. J. Carter.** 1989. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. Virology **173**:120–128.
7. **Chiorini, J. A., M. D. Weitzman, R. A. Owens, E. Urceley, B. Safer, and R. M. Kotin.** 1994. Biologically active Rep proteins of adeno-associated virus type 2 produced as fusion proteins in *Escherichia coli*. J. Virol. **68**:797–804.
8. **Gralla, J. D.** 1985. Rapid "footprinting" on supercoiled DNA. Proc. Natl. Acad. Sci. USA **82**:3078–3081.
9. **Hauswirth, W. W., and K. I. Berns.** 1979. Adeno-associated virus DNA replication: non unit length molecules. Virology **93**:57–68.
10. **Hermonat, P. L., M. A. Labow, R. Wright, K. I. Berns, and N. Muzyczka.** 1984. Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. J. Virol. **51**:329–333.
11. **Hunter, L. A., and R. J. Samulski.** 1991. Colocalization of adeno-associated virus Rep and capsid proteins in the nuclei of infected cells. J. Virol. **66**:317–324.
12. **Im, D. S., and N. Muzyczka.** 1989. Factors that bind to adeno-associated virus terminal repeats. J. Virol. **63**:3095–3104.
13. **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.
14. **Im, D. S., and N. Muzyczka.** 1992. Partial purification of adeno-associated virus Rep78, Rep68, Rep52, and Rep40 and their biochemical characterization. J. Virol. **66**:1119–1128.
15. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
16. **Labow, M. A., L. H. Graf, Jr., and K. I. Berns.** 1987. Adeno-associated virus gene expression inhibits cellular transformation by heterologous genes. Mol. Cell. Biol. **7**:1320–1325.
17. **Labow, M. A., P. L. Hermonat, and K. I. Berns.** 1986. Positive and

negative autoregulation of the adeno-associated virus type 2 genome. J. Virol. **60**:251–258.
18. **LeFebvre, R. B., S. Riva, and K. I. Berns.** 1984. Conformation takes precedence over sequence in adeno-associated virus DNA replication. Mol. Cell. Biol. **4**:1416–1419.
19. **McCarty, D. M., M. Christensen, and N. Muzyczka.** 1991. Sequences required for coordinate induction of adeno-associated virus p19 and p40 promoters by Rep protein. J. Virol. **65**:2936–2945.
20. **McCarty, D. M., T. H. Ni, and N. Muzyczka.** 1992. Analysis of mutations in adeno-associated virus Rep protein in vivo and in vitro. J. Virol. **66**:4050–4057.
21. **McCarty, D. M., J. H. Ryan, S. Zolotukhin, X. Zhou, and N. Muzyczka.** 1994. Interaction of the adeno-associated virus Rep protein with a sequence within the A palindrome of the viral terminal repeat. J. Virol. **68**:4998–5006.
22. **Mendelson, E., J. P. Trempe, and B. J. Carter.** 1986. Identification of the *trans*-acting Rep proteins of adeno-associated virus by antibodies to a synthetic oligopeptide. J. Virol. **60**:823–832.
23. **Ni, T. H., X. Zhou, D. M. McCarty, I. Zolotukhin, and N. Muzyczka.** 1994. In vitro replication of adeno-associated virus DNA. J. Virol. **68**:1128–1138.
24. **Oelze, I., K. Rittmer, and G. Sczakiel.** 1994. Adeno-associated virus type 2 *rep* gene-mediated inhibition of basal gene expression of human immunodeficiency virus type 1 involves its negative regulatory functions. J. Virol. **68**:1229–1233.
25. **Owens, R. A., M. D. Weitzman, S. R. M. Kyostio, and B. J. Carter.** 1993. Identification of a DNA-binding domain in the amino terminus of adeno-associated virus Rep proteins. J. Virol. **67**:997–1005.
26. **Rittner, K., R. Heilbronn, J. A. Kleinschmidt, and G. Sczakiel.** 1992. Adeno-associated virus type 2-mediated inhibition of human immunodeficiency virus type 1 (HIV-1) replication: involvement of p78 super(rep)/p68 super(rep) and the HIV-1 long terminal repeat. J. Gen. Virol. **73**:2977–2981.
27. **Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka.** 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc. Natl. Acad. Sci. USA **79**:2077–2081.
28. **Samulski, R. J., A. Srivastava, K. I. Berns, and N. Muzyczka.** 1983. Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. Cell **33**:135–143.
29. **Seto, E., Y. Shi, and T. Shenk.** 1991. YY1 is an initiator-sequence binding protein that directs and activates transcription in vitro. Nature (London) **354**:241–245.
30. **Shi, Y., E. Seto, L. S. Chang, and T. Shenk.** 1991. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. Cell **67**:377–388.
31. **Snyder, R. O., D. S. Im, and N. Muzyczka.** 1990. Evidence for covalent attachment of the adeno-associated virus (AAV) Rep protein to the ends of the AAV genome. J. Virol. **64**:6204–6213.
32. **Snyder, R. O., D. S. Im, T. H. 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.
33. **Snyder, R. O., R. J. Samulski, and N. Muzyczka.** 1990. In vitro resolution of covalently joined AAV chromosome ends. Cell **60**:105–113.
34. **Srivastava, A., E. W. Lusby, and K. I. Berns.** 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. J. Virol. **45**:555–564.
35. **Straus, S. E., E. D. Sebring, and J. A. Rose.** 1976. Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. Proc. Natl. Acad. Sci. USA **73**:742–746.
36. **Tratschin, J. D., J. Tal, and B. J. Carter.** 1986. Negative and positive regulation in *trans* of gene expression from adeno-associated virus vectors in mammalian cells by a viral *rep* gene product. Mol. Cell. Biol. **6**:2884–2894.
37. **Yang, Q., and J. P. Trempe.** 1993. Analysis of the terminal repeat binding abilities of mutant adeno-associated virus replication proteins. J. Virol. **67**:4442–4447.
38. **Zhou, X., and N. Muzyczka.** Unpublished data.