JOURNAL OF VIROLOGY, Feb. 1994, p. 797–804
0022-538X/94/$04.00+0

Vol. 68, No. 2

Biologically Active Rep Proteins of Adeno-Associated Virus Type 2 Produced as Fusion Proteins in *Escherichia coli*

JOHN A. CHIORINI,[1] MATTHEW D. WEITZMAN,[2] ROLAND A. OWENS,[2] ELENA URCELAY,[1] BRIAN SAFER,[1] AND ROBERT M. KOTIN[1,3]*

Molecular Hematology Branch, National Heart, Lung and Blood Institute,[1] and Laboratory of Molecular and Cellular Biology, National Institute of Diabetes and Digestive and Kidney Diseases,[2] Bethesda, Maryland 20892, and Genetic Therapy, Inc., Gaithersburg, Maryland 20878[3]

Received 29 July 1993/Accepted 25 October 1993

Four Rep proteins are encoded by the human parvovirus adeno-associated virus type 2 (AAV). The two largest proteins, Rep68 and Rep78, have been shown in vitro to perform several activities related to AAV DNA replication. The Rep78 and Rep68 proteins are likely to be involved in the targeted integration of the AAV DNA into human chromosome 19, and the full characterization of these proteins is important for exploiting this phenomenon for the use of AAV as a vector for gene therapy. To obtain sufficient quantities for facilitating the characterization of the biochemical properties of the Rep proteins, the AAV *rep* open reading frame was cloned and expressed in *Escherichia coli* as a fusion protein with maltose-binding protein (MBP). Recombinant MBP-Rep68 and MBP-Rep78 proteins displayed the following activities reported for wild-type Rep proteins when assayed in vitro: (i) binding to the AAV inverted terminal repeat (ITR), (ii) helicase activity, (iii) site-specific (terminal resolution site) endonuclease activity, (iv) binding to a sequence within the integration locus for AAV DNA on human chromosome 19, and (v) stimulation of radiolabeling of DNA containing the AAV ITR in a cell extract. These five activities have been described for wild-type Rep produced from mammalian cell extracts. Furthermore, we recharacterized the sequence requirements for Rep binding to the ITR and found that only the A and A′ regions are necessary, not the hairpin form of the ITR.

The left open reading frame (ORF) of adeno-associated virus type 2 (AAV) encodes the so-called Rep proteins. Two promoters located at map positions 5 and 19 (p5 and p19, respectively) control expression of the four proteins derived from this ORF (6, 24, 31). Processing of a common intron results in two gene products which are derived from transcripts that are initiated from either promoter (and are designated by the apparent mass of the protein in kilodaltons); Rep78 and Rep68 are produced from p5-initiated transcripts, and Rep52 and Rep40 are produced from p19-initiated transcripts (22). Plasmids containing cloned AAV yielded wild-type infectious AAV when transfected into adenovirus-infected cells (23, 28). However, mutations within the Rep gene blocked DNA replication (9).

The Rep proteins have pleiotropic effects on infected or transfected cells. Properties of the p5-initiated Rep proteins determined in vivo include the abilities to regulate AAV transcription (2, 21), activate replication of AAV (19), inhibit transcription of heterologous promoters (8, 20), and inhibit cellular transformation by bovine papilloma virus (7) or adenovirus E1a plus an activated *ras* oncogene (15). By in vitro analysis, p5-derived Rep proteins have been shown to bind to the AAV inverted terminal repeat (ITR) (11, 12, 26), have sequence- and strand-specific endonuclease activity and helicase activity (12), and bind to a defined region of human chromosome 19 at the integration locus for AAV provirus (35). These in vitro activities are thought to be essential for the viral life cycle via either lytic or latent pathways under conditions permissive or nonpermissive for DNA replication, respectively.

The effects of the Rep proteins on AAV gene expression and DNA replication are seemingly contradictory. In the absence of helper virus coinfection, the Rep proteins are inhibitors of replication and transcription (3, 19). However, in the presence of helper virus coinfection, the Rep proteins function as transactivators of expression and replication (19). The roles of the proteins in replication appear to be the result of direct interactions between Rep and the viral ITR, whereas the transcriptional effects may be mediated by yet-undetermined cellular factors.

Previous analyses of the Rep proteins were performed with proteins produced from mammalian cell extracts by use of either adenovirus- plus AAV-infected cells (11–13) or a heterologous viral expression system, e.g., the human immunodeficiency virus long terminal repeat (25, 26). The amount of Rep produced from such systems represented a small fraction of the total cellular protein; estimates of ≤1% are typical (11, 13, 26). Baculovirus expression of Rep proteins in *Spodoptera frugiperda* insect cells also resulted in yields of ≤1% of total cellular protein (26). The purification of Rep proteins from cellular extracts has been problematic. One procedure involved three serial columns: phenyl-Sepharose, DEAE-cellulose, and single-stranded-DNA–agarose (13). The partially purified Rep proteins were estimated to have been purified 200- to 1,000-fold. This extent of purification resulted in low yields of protein as a result of low levels of expression of Rep proteins and difficulties associated with the purification (11). Rep68 derived from infected-cell extracts has been purified to near homogeneity by fractionation on six serial columns, although the yield was also low (12).

The *malE* gene product of *Escherichia coli*, maltose-binding protein (MBP), has a high affinity for amylose as well as for maltose. Soluble fusion proteins generated with an MBP moiety can easily be isolated from supernatants of lysates prepared from induced *E. coli* by adsorption to and elution from an amylose resin. In this study the AAV *rep* ORF was

* Corresponding author. Mailing address: Molecular Hematology Branch, Building 10, Room 7D18, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892. Phone: (301) 496-1594. Fax: (301) 496-9985.

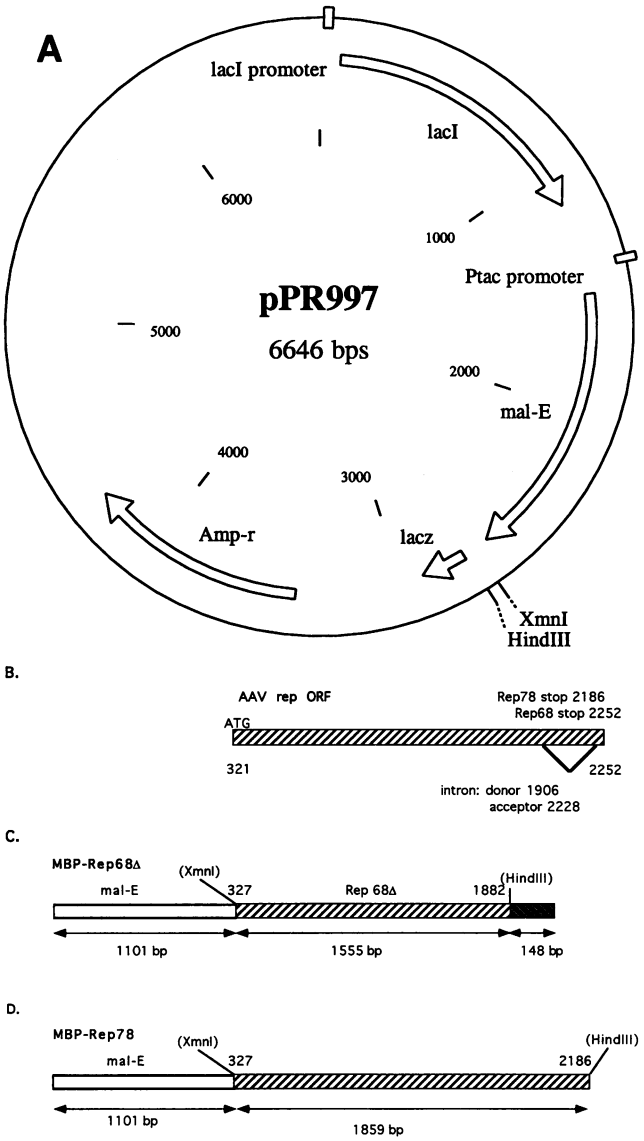

FIG. 1. (A) Map of the vector pPR997 (New England Biolabs) with the relevant restriction sites shown. (B) AAV *rep* ORF with the positions of the start codon, stop codons, and splice donor and acceptor shown. (C) pMBP-Rep68Δ construct. (D) pMBP-Rep78. The nucleotide positions are in reference to the AAV genome (30).

cloned and expressed as an MBP fusion protein in *E. coli*. Large quantities of the recombinant Rep proteins were isolated and characterized in vitro for known wild-type Rep activities. The recombinant MBP-Rep68Δ and MBP-Rep78 proteins displayed all of the activities reported for wild-type Rep proteins when assayed in vitro. Further evidence of wild-type activity was provided by the ability of MBP-Rep68Δ and MBP-Rep78 proteins to stimulate AAV replication 40- and 80-fold, respectively, in a cell extract.

MATERIALS AND METHODS

Plasmids. The coding regions for the Rep proteins (Fig. 1A) were generated by PCR amplification. A common 5′ primer corresponding to nucleotides (nt) 327 to 346 of AAV was synthesized and used for both Rep68Δ and Rep78. The Rep68Δ ORF was amplified by using a 3′ primer corresponding to the reverse complement of AAV nt 2029 to 2048. PCR amplification was performed by using cloned Pfu polymerase (Stratagene) with the buffer provided by the supplier. The PCR product was digested with *Hin*dIII, which cleaves AAV at nt 1882, and ligated into plasmid pPR997 (New England Biolabs) digested with *Xmn*I and *Hin*dIII. Plasmid pPR997 is similar to the commercially available construct pMAL-c2, the difference being that pPR997 contains the *lacI* gene whereas pMAL-c2 contains the mutant *lacI*^{q} gene (New England Biolabs). This strategy resulted in the ORF of Rep68Δ ligating in frame with the malE ORF of the vector. Ligation of the *Hin*dIII sites at the 3′ end of the *rep* ORF and the polylinker of pPR997 produced a frameshift in the *lacZ*α reading frame which resulted in the addition of 50 codons prior to the first in-frame termination codon (Fig. 1B).

A mutant MBP-Rep68Δ with a mutation in the putative nucleoside triphosphate (NTP)-binding site was produced by substitution of a *Bam*HI-*Hin*dIII fragment from the pHIVRepNTP plasmid with a lysine-to-histidine mutation at codon 340 (K340H) (26, 27). This mutation in MBP-Rep68Δ was confirmed by sequencing, and the resulting protein product was designated MBP-Rep68ΔNTP.

MBP-Rep78 was derived from MBP-Rep68Δ by extending the *rep* ORF (Fig. 1C). The additional coding sequence was generated by PCR amplification of AAV nt 1872 to 2239, which include both an overlapping region of Rep68Δ and Rep78 and the 3′ terminus of the Rep78 ORF. The natural termination codon for Rep78 was within this sequence as well. The 5′ primer corresponding to AAV nt 1872 to 1894 and a 3′ primer corresponding to the reverse complement of AAV nt 2215 to 2239 incorporated *Hin*dIII and *Xba*I sites. The PCR product was digested with *Hin*dIII and ligated into *Hin*dIII-digested pMBP-Rep68Δ. Orientation was determined by restriction analysis. The ΔITR oligonucleotides were produced synthetically. The sequences (5′→3′) are as follows: AD′ oligonucleotide, GATCAGTGATGGAGTTGGCCACTCCC TCTCTGCGCGCTCGCTCGCTCACTGAGGCCG; A′D oligonucleotide, GATCCGGCCTCAGTGAGCGAGCGAGCG CGCAGAGAGGGAGTGGCCAACTCCATCACT. The construct pΔITR was made by conversion of the 5′ overhangs of the annealed duplex to blunt ends with T4 DNA polymerase in the presence of four dNTPs. The blunt-ended annealed A′D and AD′ oligonucleotides were ligated into the *Eco*RV site of pBluescript(KS+). The sequence was confirmed by sequence analysis. The plasmid pSVori/AAV contains the AAV genome from nt 121 to 4559 and the simian virus 40 origin of replication (unpublished data) in pBluescript(KS+).

Sequence determination. The DNA sequence of the Rep78 ORF was confirmed by the dideoxynucleotide termination method (30) adapted for *Taq* polymerase by using dye terminator NTPs. Oligonucleotides (18 nt in length) for use as primers corresponded to the *rep* ORF sequence (31) at intervals of less than 300 nt. The 5′ ends of the forward primers referenced to the AAV genome are nt 321, 517, 750, 920, 1222, 1540, 1872, and 2052, and those of the reverse primers are nt 557, 844, 2295, 2535, 2797, and 2234. Polyacrylamide gel electrophoresis and acquisition of data were performed with an Applied Biosystems 373A DNA sequencer.

Protein expression. The *malE* gene is under the control of the *tac* promoter (ptac), which is repressed by the *lacI* gene product. Addition of IPTG (isopropyl-β-D-thiogalactopyranoside) prevents binding of the *lac* repressor to ptac, resulting in high levels of expression. The bacterial clones that produced a protein of the predicted molecular mass of 105 kDa, as

determined by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis, were grown on a larger scale. The fusion protein was isolated according to the protocol recommended by New England Biolabs. The bacterial pellet was resuspended in 0.05 volume of column buffer (200 mM NaCl, 20 mM Tris-Cl [pH 7.4], 1 mM EDTA, and 1 mM dithiothreitol) with 0.1 mM phenylmethylsulfonyl fluoride (PMSF) as a protease inhibitor. The bacteria were lysed by sonication for four 30-s pulses. The suspension was cleared by centrifugation at 9,000 × *g* for 20 min at 4°C. The supernatant was loaded onto a column packed with amylose resin equilibrated in column buffer plus PMSF. The column was extensively washed (e.g., >8 column volumes) with column buffer plus PMSF. The adsorbed proteins were eluted with column buffer plus PMSF and 10 mM maltose. Fractions were collected and analyzed by SDS-polyacrylamide gel electrophoresis.

Mobility shift assays. DNA-protein complexes were detected by the altered electrophoretic mobility of ^{32}P-labeled probes on nondenaturing polyacrylamide gels. The assays were performed as previously described (4, 5) with the modifications described by Jacob et al. (14). Briefly, radiolabeled probes were incubated with protein fractions at 30°C for 15 min in 25 μl. The reaction mixture contained 10 mM Tris-Cl (pH 7.5), 1 mM EDTA, 10 mM β-mercaptoethanol, 0.1% Triton X-100, 4% glycerol, and 0.5 μg of poly(dI · dC). Experiments involving competition between the radiolabeled DNA substrate and unlabeled homologous or heterologous substrate were performed as described in the figure legends.

Probes. The DNA fragments used in the mobility shift assay were prepared either by gel purification of cloned DNA fragments or by synthesis of oligonucleotides. The DNA fragments were either 3′ end labeled with [α-^{32}P]dCTP by the filling-in reaction with Klenow fragment or 5′ end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The ITR probe was produced by digesting psub201 (cloned AAV genome with *Xba*I and *Pvu*II restriction sites inboard and outboard of the ITR, respectively) (29) with the restriction enzymes *Xba*I and *Pvu*II, boiling, and quickly chilling to form the hairpin structure of the ITR. The ends of the hairpin ITR were made flush by filling in with four dNTPs and Klenow fragment. The product consists of a modified ITR plus 45 nt on the viral side, i.e., AAV nt 4490 to 4667, whereas the wild-type ITR consists of nt 4536 to 4680.

The ΔITR probes were made by 5′ end labeling either oligomer and annealing with the unlabeled complementary oligonucleotide. Integration site probe P1 was generated from pMAT50 (35). The 110-bp insert is a subfragment of the human chromosome 19 integration locus AAVS1 (16, 17). The probe was uniquely 3′ end labeled by digestion with *Hin*dIII, filling in of the overhanging 5′ end with [α-^{32}P]dCTP and Klenow fragment, and then digestion with *Eco*RI.

Helicase assay. The helicase assay was performed essentially as described by Im and Muzyczka (12, 13) except that a 17-nt primer was 5′ end labeled and annealed to M13 viral DNA (single-stranded circular template). Approximately 20 ng of this substrate was added to 20 μl of a mixture that contained 25 mM [4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid]-KOH (pH 7.5), 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.4 mM ATP, and 10 μg of bovine serum albumin per ml. The MBP-Rep proteins were added, and the reaction mixture was incubated at 37°C for 30 min. The reaction was terminated by the addition of 10 μl of 0.5% SDS–50 mM EDTA–40% glycerol–0.1% bromophenol blue–0.1% xylene cyanole FF. The reaction products were fractionated on a nondenaturing 8% polyacrylamide gel. The gel was then dried and exposed to X-ray film for autoradiography.

TRS endonuclease assay. The strand-specific endonuclease activity of MBP-Rep68Δ on the AAV ITR was detected by the methods described by Im and Muzyczka (11, 13), and the products were analyzed by electrophoresis on a 6% native TBE (Tris-borate-EDTA)-polyacrylamide gel. The reaction buffer was the same as for the helicase assay. The assay using ΔITR as a substrate was performed similarly to the ITR terminal resolution site (TRS) endonuclease assay except that approximately 20 ng of duplex ΔITR was used as the substrate in each 20 μl. The reaction was terminated by the addition of 2 volumes of stop buffer (10 mM Tris-Cl [pH 7.9], 10 mM NaCl, 0.5% SDS, 0.2 mg of yeast tRNA per ml, 20 mM EDTA, and 2 mg of proteinase K per ml). The reaction mixtures were incubated for 30 min at 37°C. The nucleic acids were extracted by phenol-chloroform and ethanol precipitated. The products were resuspended in gel loading buffer (80% formamide, 0.025% bromophenol blue, 0.025% xylene cyanole FF), denatured by heating to 80°C, and fractionated on an 8% polyacrylamide–8 M urea sequencing gel.

In vitro transcription and translation. Rep78 and Rep78 K340H were synthesized by in vitro transcription and translation as described previously (27).

In vitro replication. Cell extracts were prepared from HeLa S3 cells as described by Ward and Berns (34). The replication reactions were performed as described (34) with the following modifications: the total dCTP concentration was increased to 0.1 mM, and the reaction mixtures were preincubated for 1.5 h prior to the addition of [α-^{32}P]dCTP (specific activity, ≥5,000 Ci/mmol). The latter modification reduces the labeling from the putative repair activities to negligible levels (33). The concentration of protein in the cell extract was determined by the Bio-Rad colorimetric assay. Each assay was done with 0.1 mg of cellular protein in a final volume of 15 μl containing 1.5 μl of 10× reaction cocktail, 2 μg of creatine phosphokinase, and 10 μCi of [α-^{32}P]dCTP (5,000 Ci/mmol) (34). The plasmid DNA was double CsCl-ethidium bromide equilibrium gradient purified, and 300 ng was used in each reaction. Reaction mixtures were incubated at 34°C for 16 to 18 h. Stimulation of replication was determined by the amount of ^{32}P in the trichloroacetic acid (TCA)-precipitable material.

DNase I protection. The substrates for the DNase I protection experiments were produced from synthesized oligonucleotides corresponding to the A′D or AD′ sequences of the ITR. The oligonucleotides were individually 5′ end labeled with [γ-^{32}P]ATP and polynucleotide kinase. Equal molarities of the end-labeled oligonucleotide and the unlabeled complementary oligonucleotide were annealed prior to addition of protein. The conditions for DNase I footprinting were as described previously (11) with the following modifications: about 5×10^4 cpm of substrate was used with 0.5 μg of poly(dI-dC), and the digestion time was decreased to 30 s. The reactions were terminated as described for the ΔITR TRS endonuclease assay. The products were resuspended in gel loading buffer, denatured by heating to 80°C for 3 min, and fractionated on 6% polyacrylamide–8 M urea sequencing gels.

RESULTS

Protein purification. The recombinant MBP-Rep fusion proteins were efficiently expressed in *E. coli* following induction by IPTG. As estimated by Coomassie blue staining of an SDS-polyacrylamide gel, MBP-Rep68Δ, MBP-Rep78, and MBP-Rep68ΔNTP made up approximately 10% of the total protein in the *E. coli* lysate. The elution profile of MBP-Rep68Δ is illustrated in Fig. 2. Affinity chromatography of the lysate on amylose-Sepharose beads (New England Biolabs)

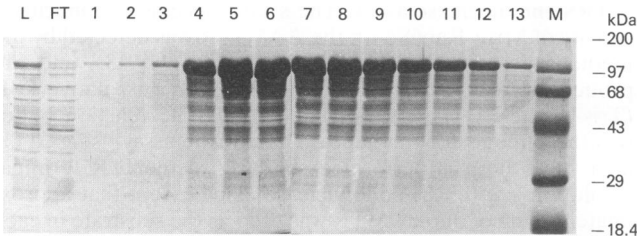

FIG. 2. Elution profile of MBP-Rep68Δ from an amylose affinity column. The supernatant from an IPTG-induced culture of *E. coli* was prepared as described in Materials and Methods. The supernatant was applied to a column containing amylose-Sepharose and washed with 10 column volumes of buffer. The adsorbed protein was eluted with column buffer containing 10 mM maltose. One-milliliter fractions were collected, and 2-μl aliquots were fractionated by electrophoresis on an SDS–8% polyacrylamide gel which was subsequently stained with Coomassie blue. Lanes: L, total *E. coli* supernatant loaded onto the column; FT, unadsorbed, flowthrough fraction; 1 to 12; aliquots of fractions eluted with 10 mM maltose; M, molecular mass standards of the sizes indicated.

quantitatively retained the 105-kDa MBP-Rep68Δ fusion protein. After extensive washing of the column with column buffer, MBP-Rep was eluted with column buffer containing 10 mM maltose. Approximately 90% of the eluted protein was full-length MBP-Rep; the remaining 10% consisted of lower-molecular-weight proteins that are likely to be early termination MBP-Rep proteins which contain part of the Rep moiety. The overall yield of MBP-Rep was typically 5 to 20 μg/ml of induced culture. The identity of the recombinant proteins was confirmed by Western blot (immunoblot) analysis with Rep-specific antiserum (32) (data not shown).

Gel shift assays. The MBP-Rep fusion proteins were tested for several activities attributed to wild-type Rep. As described previously, binding of Rep68 and Rep78 to the AAV ITR was initially done with crude protein fractions obtained from AAV-infected HeLa cells coinfected with adenovirus type 2 (1, 11) and later done with Rep expressed in 293 cells transfected with a human immunodeficiency virus long-terminal-repeat-driven construct (26). In the mobility shift assay, both of these preparations generated a tight cluster of lower-mobility bands. Similarly, MBP-Rep binding to a 5′-^{32}P-ITR hairpin probe generated a group of bands which can be inhibited by the unlabeled ITR sequence (Fig. 3A). As an additional control, expression of the MBP moiety alone generated the expected 47-kDa protein, which showed no ITR-binding activity (data not shown).

The required ITR sequence necessary for sequence-specific Rep binding was examined next. Two oligomers consisting of the A and partial D′ regions and the reverse complement were synthesized. The end-labeled annealed oligomer, designated ΔITR, was used as a probe in mobility shift assays.

Incubation of ΔITR with MBP-Rep68Δ generated a mobility shift band cluster with mobility similar to that of the cluster formed with the wild-type ITR probe (Fig. 3B). In addition, the hairpin ITR and ΔITR complexes were sensitive to direct as well as cross-competition with unlabeled probe. The data suggest, therefore, that the binding site for Rep68Δ on the ITR is largely confined to the A and A′ regions.

Gel shift assay with a human chromosome 19 DNA probe. In addition to binding the AAV ITR, wild-type Rep68 and Rep78

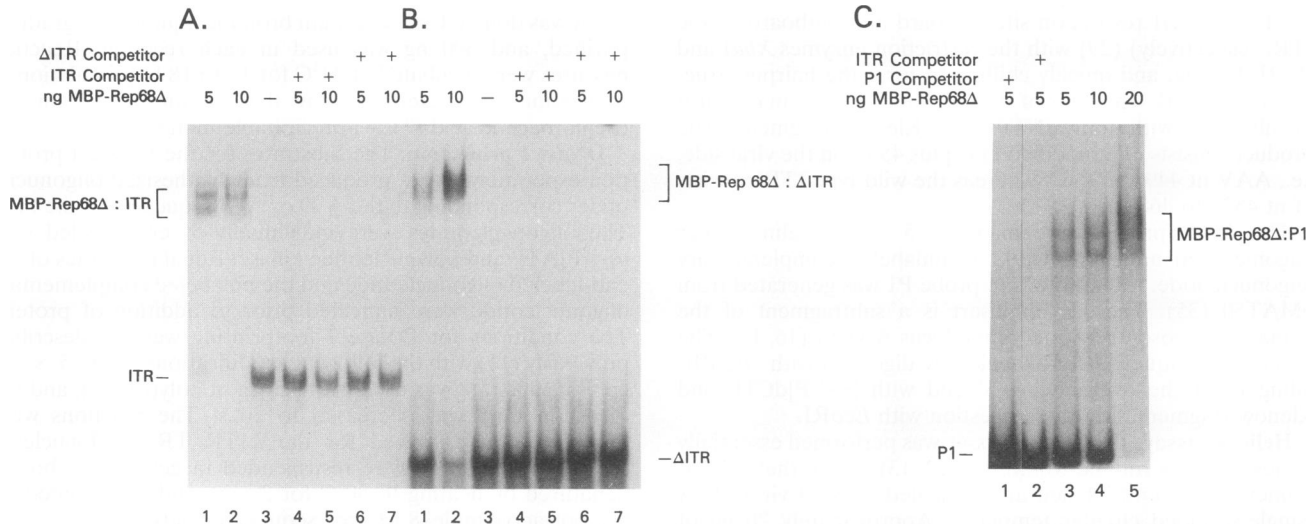

FIG. 3. Autoradiographs of the mobility shift assay. The conditions for this assay are described in Materials and Methods. Each reaction mixture contained 2 to 4 fmol of labeled probe (10,000 cpm) and 5 or 10 ng of MBP-Rep68Δ. (A) Binding of MBP-Rep68Δ to ^{32}P-labeled ITR is sensitive to competition with unlabeled ITR or ΔITR. MBP-Rep68Δ–ITR complexes were formed as described in Materials and Methods, and the positions of free ITR probe and MBP-Rep68Δ–ITR complexes are indicated by the square brackets. Lanes: 1 and 2, 5 and 10 ng of MBP-Rep68Δ, respectively; 3, ITR probe only; 4 and 5, 5 and 10 ng of MBP-Rep68Δ, respectively, with a fivefold excess of unlabeled ITR competitor; 6 and 7, same as 4 and 5 but with a fivefold excess of unlabeled ΔITR competitor. (B) Same assay as described for panel A except that the probe was ^{32}P-labeled ΔITR. The positions of the free ΔITR probe and the MBP-Rep68Δ–ITR complexes are shown. Lanes: 1 and 2, 5 and 10 ng of MBP-Rep68Δ, respectively; 3, ΔITR probe only; 4 and 5, 5 and 10 ng of MBP-Rep68Δ, respectively, with a fivefold excess of unlabeled ITR competitor; 6 and 7, same as 4 and 5 but with a fivefold excess of unlabeled ITR or ΔITR competitor. (C) Binding of MBP-Rep68Δ to the P1 sequence of the chromosome 19 AAV DNA integration site. MBP-Rep68Δ (5, 10, or 20 ng) was incubated with 4 fmol of ^{32}P-labeled P1 (12,000 cpm) as described in Materials and Methods for the mobility shift assay. The positions of free P1 probe and the MBP-Rep68Δ complexes are indicated. Both P1 and ITR competitor oligonucleotides specifically inhibited binding of the fusion proteins to the P1 probe (lanes 1 and 2). Increasing amounts of MBP-Rep68Δ (lanes 3 to 5) resulted in stoichiometric binding of the probe.

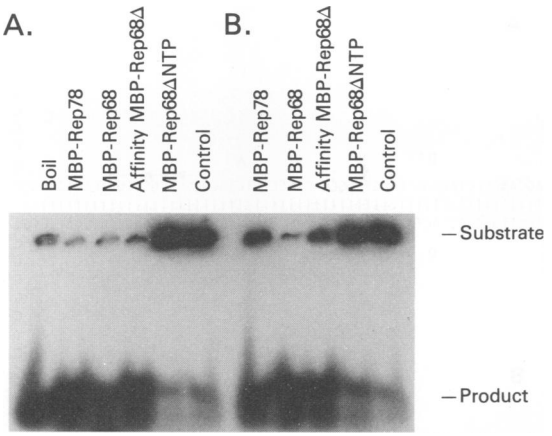

FIG. 4. Helicase assay. The recombinant MBP-Rep fusion proteins were tested for the ability to displace a labeled oligonucleotide from a circular, single-stranded DNA template (see Materials and Methods). The positions of the substrate and the free oligonucleotide product are indicated. Disassociation of the probe from the M13 viral DNA (single-stranded circular template) is the result of an ATP-dependent helicase activity previously described for AAV Rep68 and Rep78. The lanes are labeled according to the added protein. Boil, boiled substrate, no added protein; Affinity MBP-Rep68Δ, ΔITR affinity-purified Rep68Δ; Control, no added protein. Panels A and B are the same for the amount of added protein. The final volume of each added protein was adjusted to 1.0 μl by the addition of column buffer. The reaction mixtures for panels A and B contained 1 and 0.1 μg of fusion proteins, respectively.

have recently been shown to bind to a sequence of human chromosome 19 (35) identified as a preferred site for AAV DNA integration (16–18). MBP-Rep68Δ forms a stable complex with this DNA fragment (P1) containing a sequence motif also present in the Rep68 ITR-binding site (Fig. 3C). As expected, both nonradiolabeled ITR and P1 compete with the ^{32}P-P1 probe for MBP-Rep68Δ binding. Increasing the concentration of MBP-Rep68Δ to a level greater than that required to form a stable complex with all of the P1 probe does not result in the generation of any band with a lower mobility. The data suggest, therefore, that both substrates bind to a site on MBP-Rep which is mutually competed for by the P1 and ITR probes.

Helicase activity. Wild-type Rep68 and Rep78 have a DNA helicase activity that can be measured by the displacement of a labeled oligonucleotide annealed to single-stranded M13 DNA (12, 13). Both MBP-Rep68Δ and MBP-Rep78 have an ATP-dependent helicase activity as determined by this assay. Helicase assays with MBP-Rep68ΔNTP displayed approximately the same amount of displaced oligonucleotide as the reaction with no added protein (Fig. 4). Figure 4B is a repeat of the protocol of Fig. 4A, except that 1/10 of the amounts of the indicated MBP-Rep proteins was used.

TRS endonuclease assay. Wild-type Rep68 and Rep78 have a site-specific and strand-specific endonuclease activity that is critical for viral DNA replication. Cleavage at this site, the TRS, at the AD′ border region of the AAV ITR results in transfer of part of the template ITR to the daughter strand. The template strand can subsequently be repaired so that the template and daughter strands are chimeras of nascent and input DNA. The nicking or TRS endonuclease activity can be measured in vitro by using end-labeled AAV ITR as the substrate (12, 13, 27). Incubation of 5′-end-labeled AAV hairpin with MBP-Rep68Δ generated a product with a molecular weight similar to that of the control Rep78 produced by in vitro transcription and translation (Fig. 5A). Further similarities to wild-type Rep functions were demonstrated by the dominant-negative effect of the mutant MBP-Rep68ΔNTP on MBP-Rep68Δ. This effect with in vitro-transcribed and -translated wild-type Rep78 and the K340H Rep78 mutant (27) is evident in Fig. 5A.

TRS endonuclease activity of MBP-Rep68Δ and MBP-Rep78 on ΔITR substrate. Because MBP-Rep68Δ retained TRS endonuclease activity on the ITR and both MBP-Rep68Δ and MBP-Rep78 had binding affinity for ΔITR, it followed that the ΔITR may be a suitable substrate for TRS endonuclease activity for either of the fusion proteins. The results of the endonuclease assay are shown in Fig. 5B. The substrate for the reaction was the annealed, complementary oligonucleotides corresponding to the A region and 14 nt of the D′ region (and the reverse complements A′ and D) of the ITR plus a 5′ GATC overhang. The oligonucleotides were individually 5′ end labeled with γ-^{32}P and incubated with MBP-Rep68Δ or MBP-Rep78. The cleavage site can be accurately determined by comparison with the ladder for the G+A chemical sequencing reaction. The cleavage site is indicated diagrammatically in Fig. 6B and agrees with the position of the endonuclease cleavage site previously described for the intact ITR with wild-type Rep (11, 26). The other products in the reaction appear to be the result of a 3′→5′ exonuclease that either copurified with MBP-Rep68Δ or MBP-Rep78 or is an intrinsic property of these proteins. The level of endonuclease activity on the ΔITR substrate was approximately 2% of the level on the ITR substrate, suggesting that the secondary structure contributes to the efficiency of cleavage but is not essential for specific endonuclease activity.

Footprint of MBP-Rep78 on ΔITR. Specific binding of MBP-Rep78 to ΔITR was also demonstrated by DNase I footprint analysis (Fig. 7). Previously published DNase I footprints of the AAV ITR obtained by using crude preparations which contained Rep68 and Rep78 indicated specific Rep-ITR sequence interaction in the A, A′, and internal palindrome regions (11). When this analysis was performed with recombinant MBP-Rep78 and ΔITR, protection against DNase I digestion was observed only in the A and A′ regions, not in the D and D′ regions (Fig. 7).

Activities of MBP-Rep68Δ and MBP-Rep78 in an in vitro replication assay. Inserts with the AAV ITR can be excised (or rescued) from circular plasmids in vivo (23, 28) as well as in vitro. The rescue reaction in vitro is greatly stimulated by replication of the plasmid, which in previous experiments was effected by the simian virus 40 origin of replication plus exogenously added simian virus 40 large T antigen (10, 34). The effects of MBP-Rep78, MBP-Rep68Δ, and MBP-Rep68ΔNTP were tested in an in vitro replication assay with a set of plasmids to examine the effect of an intact ITR, ΔITR, and the non-ITR AAV sequences. The constructs used are pAV2 (23), which consists of wild-type AAV cloned into a pBR322 derivative; pΔITR; pSVori/AAV, which is an AAV genome lacking ITRs cloned into a vector with the simian virus 40 origin of replication; and pBluescript. Either MBP-Rep68Δ or MBP-Rep78 protein stimulated incorporation of label into TCA-precipitable material compared with the results with MBP-Rep68ΔNTP or no added MBP-Rep proteins, with pAV2 as a substrate (Fig. 8). In assays with pΔITR there was a substantial increase in incorporation of [^{32}P]dCMP into TCA-precipitable material compared with that in reactions with either no added MBP-Rep proteins or MBP-Rep68ΔNTP. The reduced incorporation of [^{32}P]dCMP with pΔITR may be attributed at least in part to the presence of a

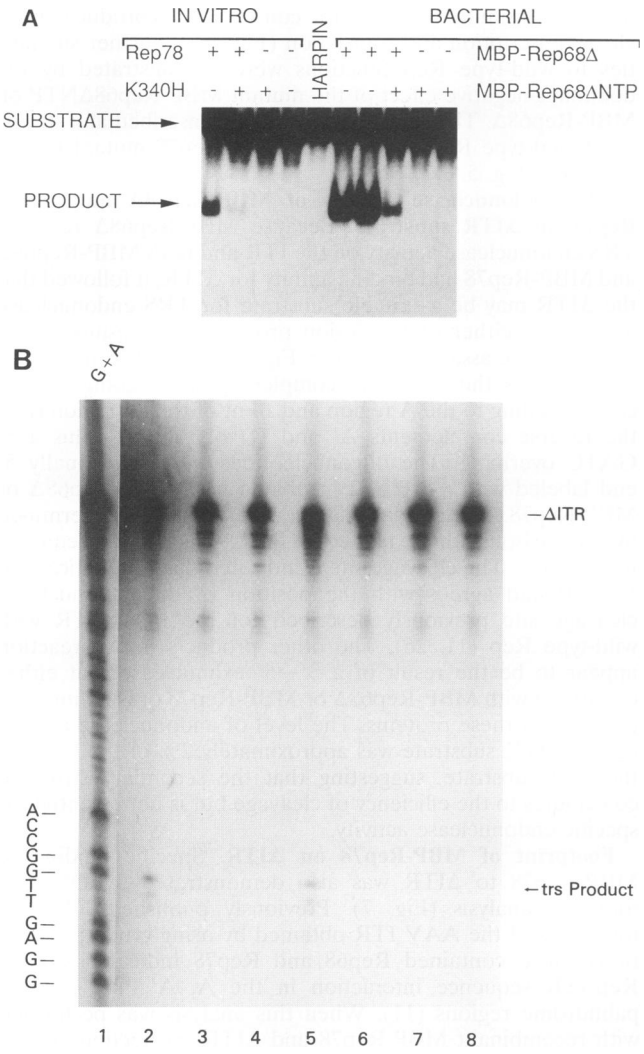

FIG. 5. TRS endonuclease assay. The TRS endonuclease assays were performed with either hairpin ITR probe or linear duplex probe generated from annealed oligonucleotides. Assays were performed as described in Materials and Methods. (A) Lanes labeled IN VITRO are TRS endonuclease reactions with 2 μl of in vitro-transcribed and -translated Rep78, Rep78 K340H, or Rep78 K340H with hairpin ITR substrate. K340H refers to the NTP binding site mutant described previously (25). The lane with a + for both Rep78 and K340H contains a mixture of approximately equivalent amounts of both proteins. Lanes labeled BACTERIAL are reactions with MBP-Rep68Δ and MBP-Rep68ΔNTP on hairpin ITR substrate. The duplicate lanes with MBP-Rep68Δ are reactions with approximately 0.5 and 0.25 μg of recombinant protein. A similar dilution was done for MBP-Rep68ΔNTP. The mixture of MBP-Rep68Δ and MBP-Rep68ΔNTP was done with approximately 0.25 μg of each recombinant protein. The position of the input hairpin ITR substrate is indicated by the arrow labeled SUBSTRATE, and the endonuclease cleavage product is indicated by the arrow labeled PRODUCT. The reactions were terminated by the addition of SDS and boiled prior to electrophoresis on a 6% polyacrylamide gel. (B) A synthetic oligonucleotide corresponding to the A and D′ sequences of the AAV ITR was 5′ end labeled and annealed to the complementary oligonucleotide (see Materials and Methods). Approximately 20 ng of duplex oligonucleotides was used as the substrate in each 20-ml reaction mixture. The products were fractionated on an 8% sequencing gel, and the gel was fixed, dried, and exposed to X-ray film for autoradiography. Lanes: 1, G+A sequencing reaction of the end-labeled oligonucleotide for use as a sizing ladder; 2, 3, and 4, 1.0, 0.1, and 0.01 μg of MBP-Rep78 added per reaction, respectively; 5, 6, and 7, 1.0, 0.1, and 0.01 μg of MBP-Rep68Δ added per reaction, respectively; 8, no MBP-Rep added. The positions of the substrate and the product are indicated.

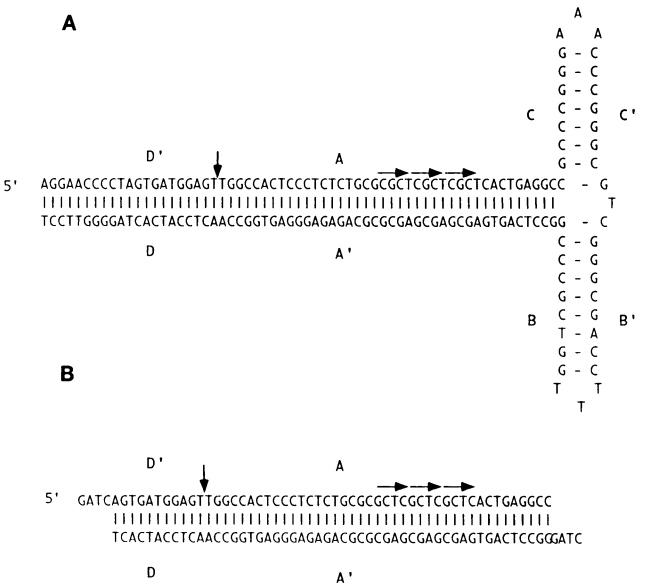

FIG. 6. (A) Sequence of the AAV ITR folded into a hairpin structure. The vertical arrow indicates the TRS, as described previously (10, 11). The three horizontal arrows are above the GCTC repeats. (B) ΔITR sequence used as a substrate for the electrophoretic mobility shift, the TRS assay, and DNase I protection assays. The vertical arrow indicates the position of the TRS as shown in Fig. 5, and the horizontal arrows are the same as in panel A.

single MBP-Rep-binding sequence and presumably a single origin of replication, compared with two origins of replication in pAV2. The stimulation of radiolabeling with MBP-Rep78 and MBP-Rep68Δ was similar for pSVori/AAV and pBluescript and was usually within a factor of two compared with the control with no added MBP-protein (Fig. 8).

The addition of MBP-Rep68ΔNTP to extracts with each of the four plasmids resulted in levels of radiolabeling similar to those observed for assays performed with no added MBP-Rep proteins. Furthermore, the degree of labeling in extracts that contained MBP-Rep68ΔNTP was uniformly low despite the presence of ITR sequences on the DNA template (Fig. 8). These results indicate that the stimulation of labeling was dependent on the addition of MBP-Rep proteins and at least the ΔITR in *cis*.

DISCUSSION

The functional characterization of AAV Rep proteins has been greatly facilitated by exploiting an *E. coli* expression system. Large amounts of Rep68Δ and Rep78 have been produced as fusion proteins with the gene product of the *E. coli malE* ORF. These fusion proteins displayed activities comparable to those of wild-type Rep68 and Rep78 expressed in mammalian systems.

The ability of Rep78 and Rep68 to bind to the viral ITR is an essential function for the viability of the virus. The fusion proteins were assayed in vitro for specific interactions with the AAV ITR. Mobility shift analysis demonstrated that both fusion proteins retained specific binding to the ITR sequences

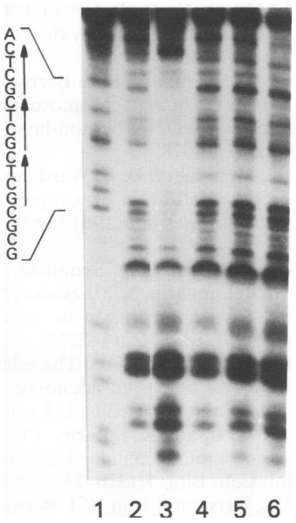

FIG. 7. DNase I protection assay of MBP-Rep78 and ΔITR sequence. The DNA substrate for the DNase I protection experiments was 5′-end-labeled synthetic A′D oligonucleotide annealed to the unlabeled AD′ oligonucleotide. Reactions conditions were as described in Materials and Methods. Lanes: 1, size markers produced by a G+A sequencing reaction of the substrate used for the footprint; 2 and 3, 2 μg of MBP-Rep78 with 30 and 90 ng of DNase I, respectively; 4, 5, and 6, 20, 40, and 60 ng of DNase I, respectively, and no MBP-Rep added. The region protected from DNase I digestion is indicated by the sequence on the left. The three arrows are parallel to the GCTC Rep-binding motif.

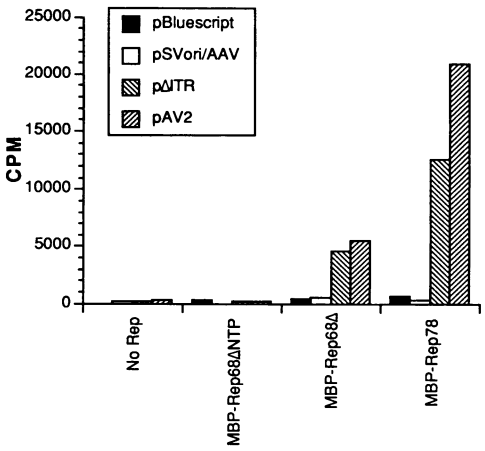

FIG. 8. Stimulation of replication in vitro. Incorporation of [^{32}P] dCMP into a TCA-insoluble fraction was measured and used to determine what effect addition of a recombinant MBP-fusion protein had on radiolabeling of DNA. Duplicate 10% aliquots of each reaction mixture were assayed. The background to be subtracted from each value was determined by the reaction with pBluescript and no Rep. A standard deviation (SD) of 0 indicates that the values determined for that set were lower than the background value. SDs were calculated for each data set with the following algorithm: $SD = \{\Sigma[x - \text{avg}(x)]^2/(n - 1)\}^{1/2}$, where x = the value of each datum point, $\text{avg}(x) = \Sigma x/n$, and n = the number of data points in a given series. q = the number of independent reactions. MBP-Rep78 and pBluescript, SD = 258.1, n = 6, q = 3; MBP-Rep78 and pSVori/AAV, SD = 89.2, n = 4, q = 2; MBP-Rep78 and pΔITR, SD = 3,660, n = 3, q = 2; MBP-Rep78 and pAV2, SD = 4,584, n = 6, q = 3; MBP-Rep68Δ and pBluescript, SD = 289.4, n = 6, q = 3; MBP-Rep68Δ and pSVori/AAV, SD = 247.2, n = 4, q = 2; MBP-Rep68Δ and pΔITR, SD = 414.2, n = 6, q = 3; MBP-Rep68Δ and pAV2, SD = 963, n = 4, q = 2; MBP-Rep68ΔNTP and pBluescript, SD = 96.5, n = 4, q = 2; MBP-Rep68ΔNTP and pSVori/AAV; SD = 0, n = 2, q = 1; MBP-Rep68ΔNTP and pΔITR, SD = 84.9, n = 4, q = 2; MBP-Rep68ΔNTP and pAV2, SD = 126, n = 4, q = 2; no Rep and pSVori/AAV, SD = 7.1, n = 2, q = 1; no Rep and pΔITR, SD = 0, n = 2, q = 1; no Rep and pAV2, SD = 266.2, n = 2, q = 1.

characteristic of the wild-type Rep68 and Rep78 proteins and that the MBP-Rep fusion proteins were about as efficient in binding to the ITR as the wild-type Rep proteins.

Production of AAV ITR substrate from cloned AAV DNA, e.g., psub201 (29) or derivatives, is an inefficient procedure, which yielded only about 40 ng of ITR per μg of plasmid. Further manipulations of the ITR, e.g., mutagenesis or ligation into concatamers for affinity purification of Rep proteins, were difficult because of the stability of the intramolecular structure. Published results of methylation interference experiments (27) indicated that the contact points between Rep68 or Rep78 and the AAV ITR are predominantly within the duplex A and A′ (A/A′) regions of the ITR, i.e., on the stem of the T-shaped structure. On the basis of these published footprint data and the recent finding that wild-type Rep78 specifically bound to the linear duplex DNA probe derived from the human chromosome 19 integration site (16, 18, 35), it seemed likely that the Rep protein binding motif was not primarily a function of DNA secondary structure. To test this hypothesis, synthetic oligonucleotides corresponding to the A/A′ regions of the ITR were used as substrates for mobility shift and endonuclease activity assays. The results presented indicated that, in contrast with earlier reports, only the A/A′ sequences are necessary for specific recognition of the Rep fusion proteins. Thus, Rep68 and Rep78 DNA binding functions are not limited to recognition of the T-shaped structure of the viral ITR.

In addition to the function of binding to either the ITR or the ΔITR, the results of a TRS endonuclease assay using ΔITR as the substrate indicate that the fusion proteins have retained the functions necessary for TRS endonuclease cleavage, e.g., sequence recognition and a strand-specific endonuclease activity. These data further indicate that the T-shaped intramolecular structure of the ITR is not necessary for proper binding or cleavage via the TRS endonuclease functions of the MBP-Rep fusion proteins. However, the TRS endonuclease activity of MBP-Rep68Δ and MBP-Rep78 on ΔITR appears to be approximately 50-fold lower than the activity on the ITR substrate, indicating that the T-shaped structure of the ITR may be required for efficient cleavage.

There is an apparent conflict between the results of the in vitro replication assay and those of the TRS endonuclease assay. The relative inefficiency of the endonuclease activity on ΔITR with respect to that on ITR does not correspond to the extent of replicated material with pΔITR relative to that with pAV2. There are two explanations for this discrepancy. First, if the ITRs in the pAV2 molecule are predominantly in the nonhairpin conformation, then the structure of the ITR would approximate that of ΔITR; i.e., both would be in linear duplex conformations, and therefore both substrates would be nicked at similar rates. Second, nicking may not be the rate-limiting step in the in vitro replication assays.

The DNA binding activity of Rep78 and Rep68 to the human chromosome 19 locus AAVS1 (35) is likely to be the determining factor in the specific integration of AAV DNA into human chromosome 19. There are no reports of a Rep$^-$ AAV construct or virus integrating into the chromosome 19 locus.

The ability of Rep68 and Rep78 to bind to linear duplex

DNA in general and to human DNA in particular raises issues concerning the more generalized effects that Rep proteins may have on cellular gene expression and DNA synthesis. For example, the presence in any context of the Rep-binding motif may be sufficient for Rep binding and in the proper context, e.g., the regulatory region of a gene, may result in either a positive or a negative transcriptional effect. Thus, the apparent cytotoxic effects of the Rep proteins are perhaps due to the direct interaction of Rep and cellular DNA.

The availability of AAV Rep proteins in large amounts will facilitate the biochemical characterization of these proteins. Fine mapping of functional domains can be done much more easily with the recombinant MBP-Rep expression system than had been possible previously. Assays to detect Rep interactions with cellular proteins are possible with larger quantities of Rep.

ACKNOWLEDGMENTS

We thank Linda Yang for expert technical assistance and Cathy Ciatto and Teresa Lumsden for their administrative skills. We thank Charlotte M. McGuinness and Sirkka R. M. Kyöstiö for critically reading the manuscript and for helpful discussions. The in vitro replication assays were possible as a result of extensive discussions with Peter Ward.

REFERENCES

1. **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.
2. **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.
3. **Berns, K. I., R. M. Kotin, and M. A. Labow.** 1988. Regulation of adeno-associated virus DNA replication. Biochim. Biophys. Acta **951**:425–429.
4. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
5. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: applications to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. **9**:3047–3060.
6. **Green, M. R., and R. G. Roeder.** 1980. Transcripts of the adeno-associated virus genome: mapping of the major RNAs. J. Virol. **36**:79–92.
7. **Hermonat, P. L.** 1989. The adeno-associated virus Rep78 gene inhibits cellular transformation induced by bovine papillomavirus. Virology **172**:253–261.
8. **Hermonat, P. L.** 1991. Inhibition of H-ras expression by the adeno-associated virus Rep78 transformation suppressor gene product. Cancer Res. **51**:3373–3377.
9. **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**:329–339.
10. **Hong, G., P. Ward, and K. I. Berns.** 1992. In vitro replication of adeno-associated virus DNA. Proc. Natl. Acad. Sci. USA **89**:4673–4677.
11. **Im, D. S., and N. Muzyczka.** 1989. Factors that bind to adeno-associated virus terminal repeats. J. Virol. **63**:3095–3104.
12. **Im, D.-S., and N. Muzyczka.** 1990. The AAV origin binding protein is an ATP-dependent site-specific endonuclease with DNA helicase activity. Cell **61**:447–457.
13. **Im, D. S., and N. Muzyczka.** 1992. Partial purification of adeno-associated virus Rep78, Rep52, and Rep40 and their biochemical characterization. J. Virol. **66**:1119–1128.
14. **Jacob, W. F., T. A. Silverman, R. B. Cohen, and B. Safer.** 1989. Identification and characterization of a novel transcription factor participating in the expression of eukaryotic initiation factor 2 α. J. Biol. Chem. **264**:20372–20384.
15. **Khleif, S. N., T. Myers, B. J. Carter, and J. P. Trempe.** 1991. Inhibition of cellular transformation by the adeno-associated virus rep gene. Virology **181**:738–741.
16. **Kotin, R. M., R. M. Linden, and K. I. Berns.** 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J. **11**:5071–5078.
17. **Kotin, R. M., J. C. Menninger, D. C. Ward, and K. I. Berns.** 1991. Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. Genomics **10**:831–834.
18. **Kotin, R. M., M. Siniscalco, R. J. Samulski, X. Zhu, L. Hunter, C. A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K. I. Berns.** 1990. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. USA **87**:2211–2215.
19. **Labow, M. A., and K. I. Berns.** 1988. The adeno-associated virus *rep* gene inhibits replication of an adeno-associated virus/simian virus 40 hybrid genome in cos-7 cells. J. Virol. **62**:1705–1712.
20. **Labow, M. A., L. J. Graf, and K. I. Berns.** 1987. Adeno-associated virus gene expression inhibits cellular transformation by heterologous genes. Mol. Cell. Biol. **7**:1320–1325.
21. **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**:515–524.
22. **Laughlin, C. A., M. W. Myers, D. L. Risin, and B. J. Carter.** 1979. Defective-interfering particles of the human parvovirus adeno-associated virus. Virology **94**:162–174.
23. **Laughlin, C. A., J. D. Tratschin, H. Coon, and B. J. Carter.** 1983. Cloning of infectious adeno-associated virus genomes in bacterial plasmids. Gene **23**:65–73.
24. **Lusby, E. W., and K. I. Berns.** 1982. Mapping of the 5′ termini of two adeno-associated virus 2 RNAs in the left half of the genome. J. Virol. **41**:518–526.
25. **Owens, R. A., and B. J. Carter.** 1992. In vitro resolution of adeno-associated virus DNA hairpin termini by wild-type Rep protein is inhibited by a dominant-negative mutant of Rep. J. Virol. **66**:1236–1240.
26. **Owens, R. A., J. P. Trempe, N. Chejanovsky, and B. J. Carter.** 1991. Adeno-associated virus rep proteins produced in insect and mammalian expression systems: wild-type and dominant-negative mutant proteins bind to the viral replication origin. Virology **184**:14–22.
27. **Owens, R. A., M. D. Weitzman, S. R. Kyöstiö, 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.
28. **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.
29. **Samulski, R. J., L. S. Chang, and T. Shenk.** 1987. A recombinant plasmid from which an infectious adeno-associated virus genome can be excised in vitro and its use to study viral replication. J. Virol. **61**:3096–3101.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
31. **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.
32. **Trempe, J. P., E. Mendelson, and B. J. Carter.** 1987. Characterization of adeno-associated virus rep proteins in human cells by antibodies raised against rep expressed in Escherichia coli. Virology **161**:18–28.
33. **Ward, P. (Cornell University Medical College).** 1993. Personal communication.
34. **Ward, P., and K. I. Berns.** 1991. In vitro rescue of an integrated hybrid adeno-associated virus/simian virus 40 genome. J. Mol. Biol. **218**:791–804.
35. **Weitzman, M. D., S. R. M. Kyöstiö, R. M. Kotin, and R. A. Owens.** Adeno-associated virus (AAV) rep proteins mediate complex formation between AAV DNA and the human integration site. Submitted for publication.