Role of the terminal repeat GAGC trimer, the major Rep78 binding site, in adeno-associated virus DNA replication

B.M. Bishop[a], A.D. Santin[a,b], J.G. Quirk[a], P.L. Hermonat[a,*]

[a] *Departments of Obstetrics and Gynecology, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA*
[b] *Division of Gynecologic Oncology, University of Brescia, Brescia, Italy*

Received 24 September 1996

Abstract The adeno-associated virus (AAV) terminal repeats (TR) are *cis* required, and the AAV encoded Rep78 protein is *trans* required, for AAV DNA replication. The Rep78 protein recognizes and interacts with at least three regions within the TR DNA. The major binding site, with the highest affinity for Rep78 binding, is within the TR stem (nt 36–16) and includes the 'core' GAGC trimer ($GAGC^3$, nt 33–22; Fig. 2) sequence. In this study mutations were made within the GAGC trimer and these mutants assayed for their ability to allow for AAV double stranded (ds DNA, prepackaging DNA replication), and single stranded DNA (ss DNA, due to virion packaging) replication. Here, it is shown that when the two inside GAGC motifs are mutated, with only motif no. 1 left intact (see Fig. 2), the resulting AAV (mutA) genome was significantly defective for both ds DNA (17% of wild type) and ss DNA (9%). If the TRs contained only the two outside motifs intact (mutB), motifs no. 1 and 2, the AAV genome had a significant but reduced level of both ds (50%) and ss (34%) DNA replication. Finally, if only the middle motif no. 2 was mutated, with motifs no. 1 and 3 left intact (mutC), the resulting DNA replication for both ds and ss forms was essentially wild type (80% that of wild type). These data suggest that the GAGC trimer plays a role in AAV DNA replication, and that GAGC motif no. 3 is the most important of the three motifs for both ds and ss DNA replication.

Key words: Adeno-associated virus; Terminal repeat; DNA replication; *cis* sequence

1. Introduction

Adeno-associated virus is a helper dependent parvovirus in which there is significant interest in its use as a gene therapy vector [1,2] and its anti-tumorigenic abilities [3–6]. As a parvovirus AAV replicates its DNA by a 'rolling hairpin' mechanism [7–9]. The terminal repeats (TR) elements, of 145 base pairs, located at each terminus of the linear AAV DNA genome, are *cis* required for AAV DNA replication [10]. Another requirement are the large AAV *rep* gene products, which are *trans* required for this AAV DNA replication [11,12]. Of the several related proteins produced from the AAV *rep* gene, the Rep68/78 proteins (the full length ORF products; Rep68 is produced by a splicing event at its 3′ end) are the most important viral proteins involved in AAV DNA replication [11,12]. Rep78 is a multifunctional protein which binds DNA, binds ATP, has DNA helicase activity, and has DNA endonuclease activity [13–20]. The Rep78 protein recognizes and interacts with at least three regions within the AAV TR DNA. The major binding site, with the highest affinity for Rep78, is within the TR stem [21–23](nt 1–41, see Fig. 2). The 'core' recognition sequence within this region is a GAGC trimer ($GAGC^3$, nt 22–33). Multimers of this repeat are preferentially bound by Rep78 from pools of randomly synthesized double stranded (ds) oligonucleotides [24]. In a second interaction, Rep78 recognizes part of the AAV TR DNA cruciform structure, or 'rabbit ear' at nt 50–54 [25], which we refer to as the minor Rep78 binding site. Mutations in these sequences result in defective DNA replication [26–28]. The major and minor Rep78 binding sites are shown within Fig. 2. Finally, Rep78 protein interacts with a sequence within the stem sequence, inboard of the GAGC trimer, and endonucleolytically nicks one strand (at nt 1), which is referred to as the terminal resolution site (*trs*) [17,18]. This nicking is essential for AAV DNA replication as it creates a 3′ primer for completing the duplication of the AAV TR sequence during DNA replication.

The DNA sequence which constitutes the major binding site for the Rep78 protein, as determined by DNAse I footprint analysis and by electrophoretic mobility shift assays (EMSA) with oligonucleotides, is a triplex concatemeric sequence of the motif 5′ GAGC 3′ (or GCTC) [16,19]. A variety of EMSA studies demonstrate that three such contiguous motifs, with appropriate adjacent sequences, allow for very strong binding by Rep78 [21–25]. Single motifs, while bound at a limited level are much weaker targets for Rep78 [21,23,29]. Two immediately adjacent motifs can often be a stronger substrate for Rep78 binding compared to single motifs [24]. Although the minor Rep78 binding site is known to be critical for *trs* activity, it has been widely assumed that the major binding site, as it is the preferential target of Rep78, likely plays a critical role in AAV DNA replication [15,16]. In this study, the role of the DNA sequence, 5′ GAGCGAGCGAGC 3′, the GAGC trimer, of the AAV terminal repeats (TR), is studied in regard to its requirement for AAV DNA replication.

2. Materials and methods

2.1. Cells

SW13 cells were maintained in Dulbecco's Modified Eagles Medium with 7% fetal bovine serum, penicillin, and streptomycin.

2.2. Construction of AAV TR mutant DNA genomes

The strategy to replace the AAV TRs with mutant sequences was carried out in a two-step process. First, the AAV genome was obtained with the TRs partially cleaved off by *Bss*HII restriction digestion of the cloned AAV plasmid pSM620. *Bss*HII cleaves within the TR sequences as shown in Fig. 1. The restricted AAV genome, minus the outer TR sequences (4.5 Kb), was isolated using Qiagen Qiaex II Gel Extraction Kit. Then, kinased synthetic oligonucleotides were ligated onto the restricted ends of the *Bss*HII AAV genome fragment to regenerate the TRs. The synthetic oligonucleotides which were used to rebuild the TRs, either as mutated or wild-type sequences, are shown in Fig. 1. A set of three oligonucleotides, TR1 (nt 30–90),

*Corresponding author. Fax: (1) (501) 686-5784.

PII S0014-5793(96)01149-0

TR5 (nt 91–107) and TR6 (nt 23–29), were used to construct the TRs. The oligonucleotide TR1 includes the AAV sequences from nt 30–90, TR5 includes nt 91–107, and TR6 includes nt 23–29. Each mutant TR was composed of TR1 plus the appropriately named TR5 and TR6 oligonucleotides. All of the TRs contained TR1, the largest and most expensive oligonucleotide, which has been used in previous studies [19,29]. The TR1 oligonucleotide was ligated with the appropriate set of TR5 and TR6 onto 300 ng of the *Bss*HII isolated 4.5 kb DNA fragment at a ratio of one AAV genomic terminus to three of each oligonucleotide to generate the mutant genomes. Ligations were *Nco*I digested, labeled by [α-^{32}P]dCTP and *E. coli* DNA polymerase Klenow fragment, and agarose gel electrophoresed to verify the extent of ligation. The transfected DNA ligations were of comparable success and averaged from 40% to 60% depending upon the specific ligation set (wt, mutA, mutB, and mutC).

2.3. Southern blot analysis of AAV DNA replication

These ligated DNAs (300 ng) were then DEAE-dextran transfected (Stratagene transfection kit) into 10 cm plates of SW13 cells at 80% confluence which were subsequently infected with Adenovirus type 2 at an MOI of 5. At 48 h post-transfection low molecular weight DNA was isolated by the method of Hirt [30], 20% of the sample agarose gel electrophoresed, Southern blotted [31], and probed with ^{32}P-labeled AAV DNA. The autoradiographs from three experiments were densitometrically analyzed using an Alpha Innotech IS-1000 documentation system and analysis software.

3. Results

After removing the resident AAV TRs by *Bss*HII restriction digestion, the TRs were replaced by the ligation of synthetic oligonucleotides (Fig. 1). These oligonucleotides resulted in the generation of either wild-type or one of three mutant AAV TRs shown in Fig. 2. Using the TR1 oligonucleotide, which we had already synthesized for experiments other than those presented here, only motifs no. 2 and 3 could be mutated. Three mutants were generated. MutA had only motif no. 1 intact. MutB had motifs no. 1 and 2 intact, mutC had motifs no. 1 and 3 intact, while wild type had motifs no. 1–3 intact. The reconstituted wild-type and mutant AAV genomes

TR1:	5′ TGAGGCCGCCCGGGCAAAGC CCGGGCGTCGGGCGACCTTT GGTCGCCCGGCCTCAGTGAG C
TR6wt	5′ GAGCGAG
TR5wt	5′ CGCGCTCGCTCGCTCAC
TR6mutA	5′ GATATCT
TR5mutA	5′ CGCGAGATATCGCTCAC
TR6mutB	5′ GAGCTCT
TR5mutB	5′ CGCGAGAGCTCGCTCAC
TR6mutC	5′ GATAGAG
TR5mutC	5′ CGCGCTCTATCGCTCAC

Fig. 1. Synthetic oligonucleotides used to construct AAV TR mutants. A set of three oligonucleotides, TR1, TR5 and TR6, were used to construct each of the mutant AAV TRs by ligating onto the *Bss*HII digest AAV genome. The oligonucleotide TR1 includes the AAV sequences from nt 30–90, TR5 includes nt 91–107, and TR6 includes nt 23–29. All of the completed TRs contained TR1, which was used in previous studies [19,29]. Each completed mutant TR was composed of TR1 plus the appropriately named TR5 and TR6 oligonucleotides.

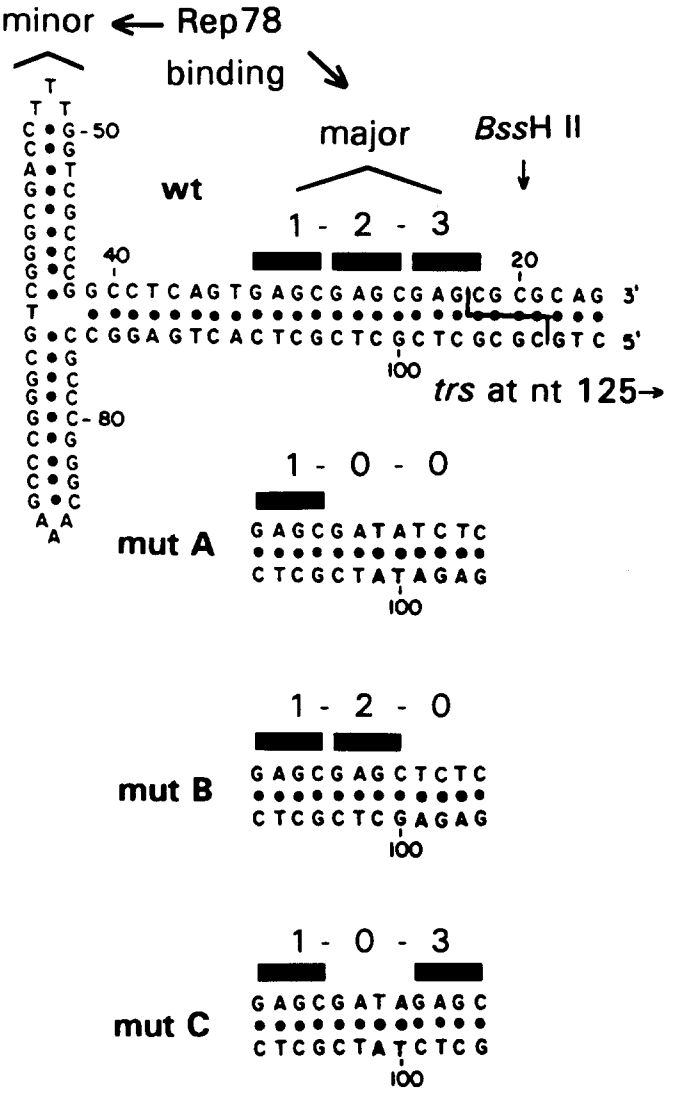

Fig. 2. Structure of wild-type and mutant AAV TRs. Shown are the structure and sequence of the wild-type and mutant TRs. These are the structures which were generated when the appropriate set of oligonucleotides (TR1, TR5, TR6) were ligated onto the *Bss*HII cut AAV genome. The complete wild-type TR is shown at the top. The major and minor Rep78 binding sites, the *Bss*HII construction site, and the *trs* site are indicated. Below the complete wt TR are shown the three mutant TR sequences but only includes the region of the GAGC trimer, in which they differ from wild type. The presence of an intact GAGC motif is indicated by a black bar.

were then transfected into Ad infected cells and AAV DNA replication observed at 48 h by Southern blot analysis (Fig. 3).

The replication levels from three experiments are quantitated in Fig. 4.

As shown in Figs. 3 and 4, the mutA with the two inside GAGC motifs mutated was the most seriously defective of the three mutants. It was defective for both double stranded (ds) (17% of wild-type) and single stranded (ss) (9%) DNA replication. The mutB which contained only the two outside motifs had less seriously reduced levels of both ds (50%) and ss (34%) DNA replication. The third mutant, mutC, with only the internal motif 2 mutated, was essentially wild type for both ds (80%) and ss (80%) DNA replication. From these data it is clear that mutating the GAGC motifs had effects on both ds

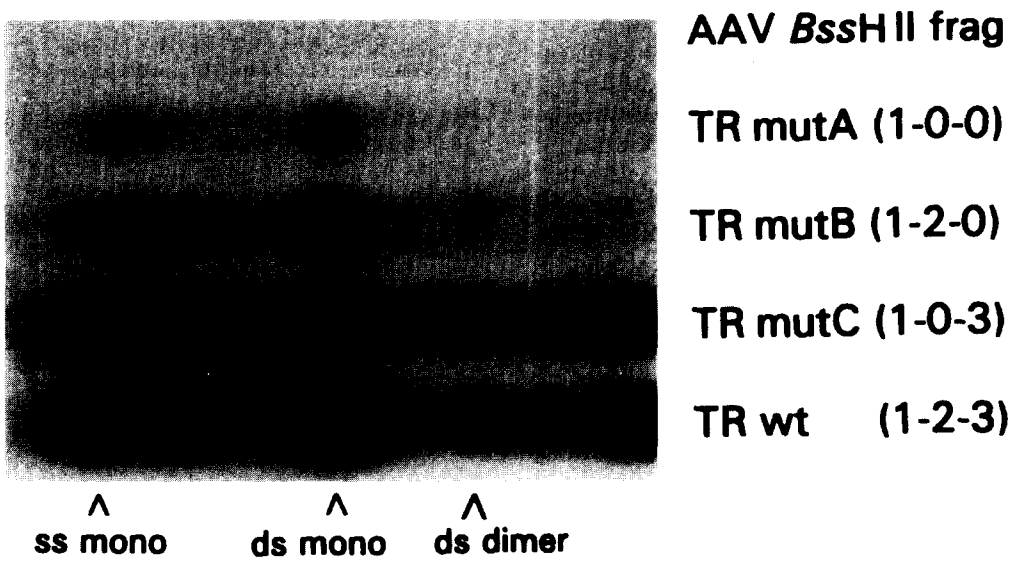

Fig. 3. Replication of AAV TR mutants by Southern blot analysis. 1/3 μg of the indicated AAV genome were transfected SW13 cells by DEAE-Dextran transfection which were subsequently infected with adenovirus. After 36–48 h low molecular weight DNA was isolated by Hirt DNA extraction. Isolated DNA was then agarose gel electrophoresed, Southern blotted, and probed with ^{32}P-labeled AAV DNA, then autoradiographed. Note that wild-type AAV replicates at the highest level, followed by mutC, mutB, and finally mutA. The positions of double stranded (ds) and single stranded (ss) replication products are indicated. Ds monomer DNA is believed to be the mature pre-packaging product, while ss DNA is DNA isolated from virus particles (complete or incomplete). AAV virions contain + or − ss DNA strands.

and ss AAV DNA replication. However, the effects were not as great as anticipated by results from the in vitro electrophoretic mobility shift assays measuring Rep78s affinity for various GAGC related DNAs [21–25,29].

4. Discussion

This study demonstrates that the complete GAGC trimer is not essential for a near-wild type AAV DNA replication phenotype. These data suggest that the in vitro studies demonstrating Rep78's affinity for various GAGC multimers and structures are not good models in anticipating and analyzing Rep78's role in DNA replication. Leaving only motif no. 1 intact resulted in the most defective replication phenotype of the three TR mutants. The inclusion of motif no. 2 in addition to no. 1 resulted in a limited improvement, but the inclusion of motif no. 3 with no. 1 resulted in a close to wild-type phenotype. We interpret these data as suggesting that the inboard GAGC motif no. 3, closest to the *trs* site, is the most important of the three motifs for both AAV ds and ss DNA replication. We have not formally observed the ability of Rep78 to bind these mutant DNAs. However, it is known from multiple previous studies that these mutant sequences will be significantly reduced in Rep78 binding affinity compared to the fully wild type sequences as any mutations within these sequences strongly inhibit the Rep78-DNA interaction [20,21,24,29]. One study in particular demonstrated that the central GAGC motif was most critical for Rep78 binding [25]. However, our mutant TR mutC, mutated in this critical site, was essentially wild type in its replication abilities. These data indicate that the straightforward ability and affinity of Rep78 to recognize and bind the major Rep78 binding site of the TR DNA is not the most important interaction for AAV DNA replication. It should be noted that one previously analyzed mutant within the minor binding site, mutant 1205, has been described previously as replication defective [10]. However, upon observing Figure 3 of this paper [10] this mutant may be replicating at some level, generating a very high molecular weight form of AAV DNA (described as form II in the figure). These data, taken together, are consistent with the hypothesis that Rep78's interaction with the minor binding site at the end of the rabbit ear, and subsequent nicking at the *trs* site, are the most important Rep78-DNA interactions for AAV DNA replication [21,23]. Furthermore, all these data suggest that these two latter interactions appear to be linked.

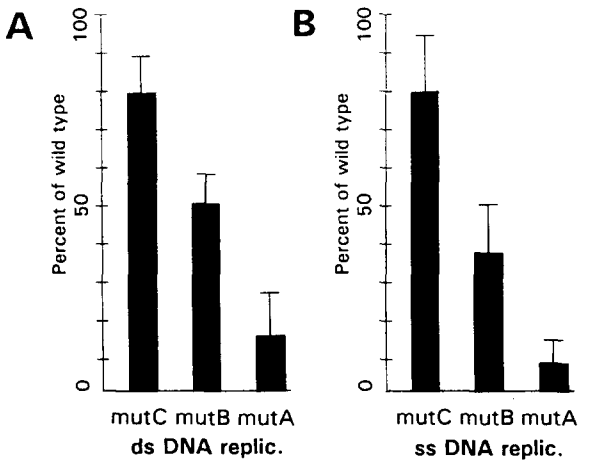

Fig. 4. Quantification of AAV mutant ds and ssDNA replication. The autoradiographs from three experiments were densitometrically analyzed using an Alpha Innotech IS-1000 documentation system and analysis software. Shown are the median and standard deviation of the three experiments. (A) Analysis of the ds DNA band. (B) Analysis of the ss DNA band.

Another possible function of the major Rep78 binding site is as a packaging signal for the AAV genomic DNA. It is known that Rep78 interacts with the capsid proteins in large complexes [32,33]. Thus, it could be hypothesized that Rep78

is involved in DNA packaging and virion maturation. The generation and demonstration of ss DNA in the Southern blots is believed to be due to the packaging, or sequestering, of these ssDNA strands into virus particles (AAV being a ssDNA virus) [11,12]. Thus, if the GAGC trimer is the packaging signal, then mutations within this sequence should result in greatly reduced ss DNA levels. However, our data demonstrating that mutC accumulates almost as much ss DNA as does wild type is inconsistent with the GAGC trimer being the packaging signal. These data leave us in quandary as to what is the specific role of the GAGC trimer. As it is clearly the site within the AAV TR where Rep78 binds most strongly this would seem to imply an important role for this interaction. One possibility, not addressed in this study but raised by others [34], is that the Rep78-major binding site interaction may be largely responsible for the site specific integration of AAV into human chromosome 19. The AAV TR mutants which we have generated here should be very useful in studies which investigate this hypothesis.

References

[1] Hermonat, P.L. and Muzyczka, N. (1984) Proc. Natl. Acad. Sci. USA 81, 6466–6470.
[2] LaFace, D., Hermonat, P.L., Wakeland, E. and Peck, A. (1988) Virology 162, 483–486.
[3] Ostrove, J.M., Duckworth, D.H. and Berns, K.I. (1981) Virology 113, 521–533.
[4] Hermonat, P.L. (1989) Virology 172, 253–261.
[5] Hermonat, P.L. (1991) Cancer Res. 51, 3373–3377.
[6] Hermonat, P.L. (1994) Cancer Res. 54, 2278–2281.
[7] Muzyczka, N. (1994) Curr. Top. Microbiol. Immunol. 158, 98–129.
[8] Hong, G., Ward, P. and Berns, K.I. (1992) Proc. Natl. Acad. Sci. USA 89, 4673–4677.
[9] Hong, G., Ward, P. and Berns, K.I. (1994) J. Virol. 68, 2011–2015.
[10] Samulski, R.J., Srivastava, A., Berns, K.I. and Muzyczka, N. (1983) Cell 33, 135–143.
[11] Hermonat, P.L., Labow, M.A., Wright, R., Berns, K.I. and Muzyczka, N. (1984) J. Virol. 51, 329–340.
[12] Tratschin, J.D., Miller, I.L. and Carter, B.J. (1984) J. Virol. 51, 611–619.
[13] Ashktorab, H. and Srivastava, A. (1989) J. Virol. 63, 3034–3039.
[14] Im, D.S. and Muzyczka, N. (1989) J. Virol. 63, 3095–3104.
[15] Im, D.S. and Muzyczka, N. (1992) J. Virol. 66, 1119–1128.
[16] Im, D.S. and Muzyczka, N. (1990) Cell 61, 447–457.
[17] Snyder, R.O., Samulski, R.J. and Muzyczka, N. (1990) Cell 60, 105–113.
[18] Chiorini, J.A., Wetzman, M.D., Owens, R.A., Urcelay, E., Safer, B. and Kotin, R.M. (1994) J. Virol. 68, 797–804.
[19] Batchu, R.B., Miles, D.A., Rechtin, T.M., Drake, R.R. and Hermonat, P.L. (1995) Biochem. Bioph. Res. Commun. 208, 714–720.
[20] Wonderling, R.S., Kyostio, S.R. and Owens, R.A. (1994) J. Virol. 69, 3542–3548.
[21] McCarty, D.M., Pereira, D.J., Zolotukhin, S., Zhou, X., Ryan, J.H. and Muzyczka, N. (1994) J. Virol. 68, 4988–4997.
[22] Snyder, R.O., Im, D.S., Ni, T., Xiao, X., Samulski, R.J. and Muzyczka, N. (1993) J. Virol. 67, 6096–6104.
[23] Chiorini, J.A., Wiener, S.M., Owens, S.R., Kyostio, R.M., Kotin, R.M. and Safer, B. (1994) J. Virol. 68, 7448–7457.
[24] Chiorini, J.A., Yang, L., Safer, B. and Kotin, R.M. (1995) J. Virol. 69, 7334–7338.
[25] Ryan, J.H., Zoloti, S., Muzyczka, N. (1996) J. Virol. 70, 1542–1553.
[26] Bohensky, R.A. and Berns, K.I. (1989) J. Mol. Biol. 206, 91–100.
[27] Bohensky, R.A., Lefebvre, R.B. and Berns, K.I. (1988) J. Virol. 166, 316–327.
[28] Lefebvre, R.B., Riva, S. and Berns, K.I. (1984) Mol. Cell. Biol. 4, 1416–1419.
[29] Batchu, R.B., Kotin, R.M. and Hermonat, P.L. (1994) Cancer Lett. 86, 23–31
[30] Hirt, B. (1967) J. Mol. Biol. 26, 365–369.
[31] Southern, E.M. (1975) J. Mol. Biol. 98, 503–518.
[32] Wistuba, A., Weger, S., Kern, A. and Kleinschmidt, J.A. (1995) J. Virol. 69, 5311–5319.
[33] Prasad, K.-M. and Trempe, J.P. (1995) Virology 214, 360–370.
[34] Weitzman, M.D., Kyostio, S.R., Kotin, R.M. and Owens, R.A. (1995) Proc. Natl. Acad. Sci. USA 91, 5808–5812.