JOURNAL OF VIROLOGY, June 1995, p. 3789–3798 Vol. 69, No. 6
0022-538X/95/$04.00+0

E1 Recognition Sequences in the Bovine Papillomavirus Type 1 Origin of DNA Replication: Interaction between Half Sites of the Inverted Repeats

ROBERT MENDOZA,† LEENA GANDHI, AND MICHAEL R. BOTCHAN*

Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, University of California, Berkeley, Berkeley, California 94720-3204

Received 16 December 1994/Accepted 14 March 1995

The E1 protein encoded by bovine papillomavirus type 1 (BPV-1) is required for viral DNA replication, and it binds site specifically to an A/T-rich palindromic sequence within the viral origin of replication. The protein is targeted to this site through cooperative interactions and binding with the virus-encoded E2 protein. To explore the nature of the E1 binding site, we inserted a series of homologous DNA linkers at the center of dyad symmetry within the E1 recognition palindrome. The effects of these modifications indicated that the E1 recognition palindrome can be separated into functional half sites. The series of insertions manifest a phasing relationship with respect to the wild-type BPV-1 genome in that greater biological activity was measured when full integral turns of the DNA helix separated the palindrome than when the separations were half-turns. This phasing pattern of activity was observed to occur in a variety of biological phenotypes, including transformation efficiency, stable plasmid copy number in cell lines established from pooled foci, and transient replication of full-length viral genomes. For replication reporter constructs where E1 and E2 are supplied in *trans* by the respective expression vectors, distance between the half sites seems to play a major role, yet the phasing relationships are measurable. DNase I protection studies showed that E1 bound very poorly to the construct containing a 5-bp linker, and binding was close to the wild-type level for the 10-bp insertion, consistent with a requirement for a phasing function between half sites with a modulus of 10 bp. Binding to the 15- and 20-bp insertion mutants was weak, but only for the 20-bp insertions was protection over both halves of the palindrome measurable. As it has been previously reported that the 18-bp palindrome contains sufficient nucleotide sequence information for E1 binding, we speculate that a minimal E1 recognition motif is presented in each half site. A comparison between this sequence and that of an upstream region that also binds E1 (the E2RE1 region) revealed a common pentanucleotide motif of APyAAPy. Mutants with substitutions of the ATAAT elements within E2RE1 failed to bind E1 protein. We present models for how repeats of the pentanucleotide sequence may coordinate E1 binding at the dyad symmetry axis of the origin and compare the DNA sequence organization of BPV-1 with those of the simian virus 40 and polyomaviruses at their origins of DNA replication.

In vivo and in vitro DNA replication studies of bovine papillomavirus type 1 (BPV-1) have demonstrated that the virus-encoded E1 and E2 proteins are important for replication of the viral DNA (40, 45). The BPV-1 E1 protein, which is likely to be a good model for most other papillomavirus E1 proteins, is an ATP-dependent DNA helicase that binds to an A/T-rich palindromic sequence at the viral origin of DNA replication (33, 40, 42, 45, 46). The E2 protein helps assemble the E1 protein by cooperatively binding with it on the origin site, where both E1 and E2 interact by protein-protein contacts and where both proteins touch specific DNA recognition motifs. The E2 protein thus serves as an initiator protein by increasing the affinity of the preinitiation complex for DNA and by increasing the specificity of E1 for origin sequences (25, 33, 34, 35a, 38, 46, 47). With an ATP nucleotide, E1 can induce distortions in the DNA (8) which are assumed to be present in precursor complexes for the next steps of the initiation event. The high-affinity E1:E2:DNA ternary complex may also be critical for replication initiation in vivo, in contrast to the lower-affinity E1:DNA complex, competing with histone octamers for the origin DNA sequences and helping in the assembly of cellular factors required for subsequent initiation events (18). The fate of E2 in subsequent steps in the initiation process is unclear, but some indirect and speculative arguments have been made which are consistent with a model wherein E2 leaves the duplex before significant unwinding occurs (20, 21). In any case, E1, along with the cellular RPA, can unwind the duplex and initiate DNA synthesis in concert with the cellular DNA polymerase:alpha primase (15, 17, 26, 27).

More detailed information about the nature of the E1 and E2 complexes which form on the origin site are required to unravel the sequence of steps that lead to the initiation of DNA replication. This is a difficult question to attack definitively, because complexes may form that are not necessarily direct precursors to the initiation reaction but represent either regulatory structures or partial complexes which may be unstable because of the absence of other factors or simply in vitro artifacts. In view of these issues, we point out that the number of E1 monomers and E2 dimers required for the initiation of replication has not been determined, and though we understand in great detail the nature of the E2

* Corresponding author. Mailing address: Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, 401 Barker Hall, University of California, Berkeley, Berkeley, CA 94720-3204. Phone: (510) 642-7057. Fax: (510) 642-7000.

† Present address: Center for Molecular Genetics, University of California, San Diego, La Jolla, CA 92093.

DNA recognition motifs (9, 10, 19, 24), little is known about the E1 recognition sequence. A dissection of the E1 recognition motif should thus provide information relevant to how the helicase binds to the origin site.

In this study, we address the significance of the inverted repeat sequences that are evident at the E1 DNA binding site of the origin center. The inverted repeats which define the core of the replication origin can be drawn in various ways, and the sequence on each side of the center of dyad symmetry has internal substructure (e.g., repeats). A previous study concluded that an 18-bp sequence that encompasses the center of dyad symmetry was sufficient and perhaps necessary for E1 binding (13). Our hypothesis was that the 18-bp sequence contains the binding sites for several E1 molecules, wherein each side of the palindrome binds an integral number of E1 monomers. The 18-bp sequence was thus thought to include an array of E1 binding sites that contributed to the cooperative assembly of a specific structure suitable for initiating bidirectional replication (32, 43, 45). This notion was influenced by the following facts. (i) E1 has structural and functional homologies to the simian virus 40 (SV40) T antigen with regard to domain organization (e.g., amino acid distances between the domains of the proteins that bind DNA and the domains of the proteins that bind ATP) (3, 22, 27, 38). Large T antigen is known to bind to a pentameric GAGGC sequence and to organize itself into two oligomeric hexamers on each side of the SV40 origin of DNA replication (5, 6, 23, 28). Given the structural similarities between E1 and T antigen, it was reasonable to search for pentameric sequences which might define the E1 binding motif within the palindromic DNA sequences which define the origin of DNA replication. (ii) E1 does indeed unwind DNA bidirectionally (7a, 19), and a twofold symmetric complex which could drive this unwinding would be a simple way to physically achieve the unwinding, whether or not the helicase machines wind through DNA in opposite directions or, alternatively, if the DNA is spooled through the helicase complex and comes out denatured through some internal port (41, 44). (iii) We had previously shown that E1 can also bind specifically and footprint to a region immediately upstream of the center of dyad symmetry at the origin core. This E1 binding region is found between BPV-1 E2 binding sites 8 and 9 in the upstream regulatory region (URR) (25, 45, 47). While this region does not contain palindromic sequences as does the origin site, small pentanucleotide repeats are found in common with the origin sequences. In order to test the simple idea that the dyad symmetry sequences at the origin consist of two half sites that might contain interacting E1 molecules and that these sites contain DNA binding motifs homologous to the upstream sites, we created mutants altered at both sites and carried out a series of binding and functional studies of the mutant and wild-type DNAs.

MATERIALS AND METHODS

Plasmid constructions. BPV genomes harboring linker insertions within the origin of replication were manufactured by using the plasmid pSS3, which contains the entire BPV-1 genome, linearized with *Bam*HI and inserted into the *Bam*HI site of the pUC 18 polylinker (43). The homologous set of 5-, 10-, 15-, and 20-bp linker insertions were manufactured by annealing complementary synthetic oligonucleotides and ligating the linkers into the unique *Hpa*I restriction site (nucleotide [nt] 3) of the BPV-1 genome (see Fig. 1 for sequences). These genomes were designated BPV L.I. 5C, BPV L.I. 10C, BPV L.I. 15C, and BPV L.I. 20C. The plasmid pKSO contains BPV origin sequences from nt 7805 to 100 cloned into the polylinker of pBluescript KS^+ (Stratagene) at the *Bam*HI and *Eco*RI restriction sites (45). The plasmids pKSO L.I. 5C, pKSO L.I. 10C, pKSO L.I. 15C, and pKSO L.I. 20C harbor the linkers indicated and were manufactured by PCR DNA amplification with the oligonucleotide primers B100 (5′-CGAATTCAGGTCCATGTGA-3′) and D7805 (5′-AGGATCCTTAAAC TACAGAC-3′) and the corresponding BPV genome as a template. The amplified products were digested with *Bam*HI and *Eco*RI, and the approximately 240-bp fragments were ligated into the polylinker of pKS^+ digested with the same enzymes. All plasmids containing linker insertions were sequenced by using the Sequenase kit and protocols supplied by U.S. Biochemicals.

Substitution mutations within the E2RE1 region were constructed by using PCR mutagenesis to alter 15-bp stretches of DNA sequence. Restriction sites in the context of the altered sequence were inserted into the substituted areas with the effect of changing most purines to pyrimidines and vice versa. In this way, spacing of flanking viral sequences was maintained while specific sites were altered. The P1 construct contains the altered sequence 5′-CACGGCGTGGTC CGT-3′ in place of the DNA at the region from nt 7684 to 7698 and was made by using two sets of primers to amplify BPV sequences from nt 7349 to 7698 (*Mlu*I-*Pfl*MI) and nt 7684 to 619 (*Pfl*MI-*Eag*I) from the plasmid pSS3. The primer at position 7349 (*Mlu*I site) contained only BPV-1 sequence, whereas the primer at position 7698 (3′-CGACCCTACTGTCGAAGGTGCCGCACCAG GCAGCG-5′) had BPV sequence to the 3′ side of the underlined nucleotides. The underlined nucleotides contain within them sequences for a *Pfl*MI restriction site. Another primer (5′-CGCCCACGGCGTGGTCCGTGGCGCATAAT CAGC-3′) containing a *Pfl*MI site complementary to the one shown above was used to amplify sequence from nt 7684 to 619 (*Eag*I) with a wild-type primer at the *Eag*I site. After amplification of both segments, the PCR products were hydrolyzed with *Pfl*MI and *Mlu*I or *Pfl*MI and *Eag*I. Upon ligation at the complementary *Pfl*MI restriction site ends, the fragment was inserted directly into pSS3 cut with *Eag*I and *Mlu*I. The construct S2 contains altered sequence, 5′-CGTCGACTAGGCCGG-3′, at nt 7704 to 7718 and was constructed in a fashion similar to that of P1. To make S2, a *Sal*I site was constructed at the junction of amplified fragments from nt 7349 to 7718 and nt 7704 to 619. These fragments were ligated and reinserted into the *Eag*I-*Mlu*I fragment of pSS3. In this case, the oligonucleotide primers used to alter the sequence were 5′-CGCGTCGACTAGGCCGGGGTGAGGACAAGCTAC-3′ (nt 7704 to 619) and 3′-TATAACTTACCCGCGGCAGCTGATCCGGCCGCC-5′ (nt 7349 to 7718). Both substitutions were confirmed by sequencing and restriction digest analysis.

Expression vectors. The pCG expression vector employs the cytomegalovirus promoter for gene expression and is described by Tanaka and Herr (36). Derivatives of the pCG vectors that express BPV-1 proteins were generously provided by A. Stenlund (Cold Spring Harbor Laboratory). The vector pCG-Eag expresses the BPV E1 protein and contains BPV sequences extending from nt 619 to 4450 with a deletion of nt 2766 to 3881. The vector pCG-E2 expresses the BPV E2 protein and contains BPV sequences from nt 2602 to 4450 with a point mutation at nt 3092 that changes the initiation codon for E2C to an isoleucine codon.

Cells. C127 (7) cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS), 100 μg of streptomycin per ml, and 100 U of penicillin per ml. VI216 is a BPV-transformed cell line that harbors approximately 150 episomal copies of the viral genome (37).

Transient replication assays. Transient replication assays were performed essentially as described by Ustav and Stenlund (39), with some changes. C127 cells were split 24 h prior to electroporation and seeded at a density of 7×10^5 cells per 10-cm-diameter dish. After 24 h, the cells were washed with phosphate-buffered saline and trypsinized with 2 ml of trypsin per dish. The cells were pooled and transferred to a 50-ml tube (Falcon) on ice containing 5 ml of DMEM with 10% FCS and 5 mM BES (*N*, *N*-bis[2-hydroxyethyl]-2-aminoethanesulfonic acid, pH 7.20). The cells were pelleted in a clinical centrifuge and resuspended to a concentration of 2×10^7 cells per ml in DMEM plus 10% FCS and 5 mM BES. The DNA sample, as indicated in the figure legends, and 50 μg of sheared salmon sperm DNA were mixed together with 0.25 ml of the cell suspension in a Microfuge tube (Eppendorf) and transferred to an electroporation cuvette. The cuvette (0.4-cm gap) was charged with 270 V at a capacitance of 960 μF by using a Bio-Rad Gene Pulser. The cell material was removed from the cuvette and directly transferred to 100 ml of DMEM with 10% FCS. Typically, the cells were plated on 10 10-cm-diameter dishes. A Hirt extraction of one plate of cells was performed for each transient replication time point.

Preparation of DNA for electroporation. DNA samples to be electroporated were first linearized. The BPV genomes were released from the pUC 18 vector sequences by digestion with *Bam*HI. The E1 and E2 expression vectors were typically linearized with either *Xmn*I or *Xho*I restriction endonuclease. The linearized DNA was purified by extraction with phenol-chloroform and chloroform. The DNA was precipitated, washed with 70% ethanol, and resuspended in 1 mM Tris-Cl–0.1 mM EDTA, pH 8.0.

Hirt extraction. Low-molecular-weight DNA was extracted by the protocol of Hirt (11). Briefly, cells were lysed with 800 μl of Hirt lysis buffer (10 mM Tris-Cl, 10 mM EDTA, 0.6% sodium dodecyl sulfate [pH 7.7]). The lysate was collected in a 2-ml microcentrifuge tube to which was added 200 μl of 5 M NaCl. The tubes were inverted several times and placed on ice for 30 min. The lysate was spun for 10 min in a microcentrifuge at 4°C, and the supernatant was transferred to a new tube. The supernatant was vortexed after the addition of 400 μl of phenol and again after the addition of 400 μl of chloroform. The lysate was spun in a microcentrifuge for 6 min, and 850 μl of supernatant was removed and transferred to a new tube. A 750-μl volume of chloroform was added, and the solution was vortexed and then given a 6-min spin in a microcentrifuge. A 700-μl volume of the supernatant was removed, and the DNA was precipitated by addition of 42 μl of 5 M NaCl and 660 μl of isopropanol. Samples were stored at −20°C until

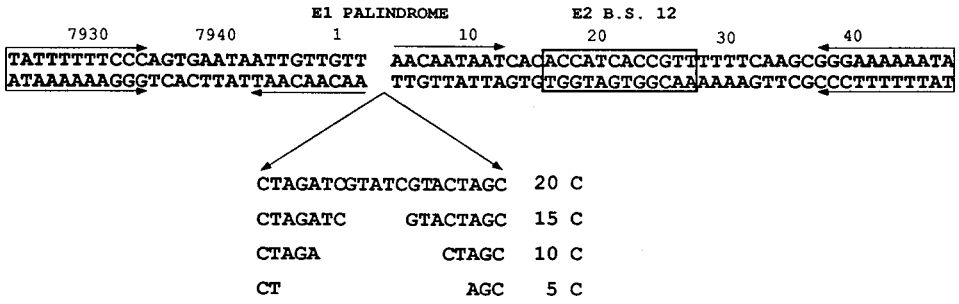

FIG. 1. DNA sequences of the BPV-1 genome which encode the replication origin. The palindromic sequence known to be important for E1 binding, the E2 binding site (B.S.) 12, and flanking A/T-rich sequences are also indicated. A homologous series of linkers labelled 20C, 15C, 10C, and 5C were inserted as shown at the center of the dyad axis at nt 3. See Materials and Methods for details of the construction.

samples from all of the time points were collected for the replication assay. The DNA was pelleted by spinning for 10 min in a microcentrifuge. The DNA pellet was washed with 80% ethanol and spun again for 10 min. The pellet was resuspended in 20 μl of 10 mM Tris-Cl–1 mM EDTA (pH 8.0). RNase A was added to 20 mg/ml, and the samples were incubated at 65°C for 20 min. A 10-μl volume of restriction endonuclease mix (6 U of *Dpn*I and 6 U of *Xba*I) was added to each sample to linearize the replicated DNA and digest the unreplicated DNA. Restriction digests were incubated for 10 h at 37°C. DNA samples were subjected to gel electrophoresis on 0.8% agarose gels in Tris-borate-EDTA (TBE) buffer. Viral DNA was detected by Southern blot analysis with nick-translated BPV-1 DNA.

Footprint analysis. DNase I protection analyses were performed as previously described (14, 45). Competitor DNA was not included in the reaction mixtures. The wild-type pKSO plasmid and derivatives containing linker insertions were digested with *Eco*RI (nt 100) and labeled for 1 h at 37°C with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase (Boehringer Mannheim) and kinase buffer (50 mM Tris-Cl [pH 7.7], 7 mM $MgCl_2$, 1 mM dithiothreitol). The labeled DNA was digested with *Bam*HI to generate the 240-bp fragment used in the footprint analyses. For the generation of labeled templates for reactions involving the wild-type E2RE1 region and the P1 and S2 mutants, the respective constructs in the context of pSS3 were digested with *Apa*LI and *Bgl*I and labeled for 1 h at 37°C with [γ-^{32}P]ATP as noted above. The labeled DNA was digested with *Ppu*MI, which releases any label incorporated at the *Bgl*I (3′) end. The E1 and E2 proteins were purified as described by Yang et al. (45).

Cell transformation. C127 cells were transfected via electroporation with linear viral DNA as noted above. Cells were plated on 10-cm-diameter dishes and cultured for 2 weeks prior to formalin fixation and staining with methylene blue. Stable cell lines were established from pooled foci by culturing the transformed cells through 10 passages (approximately 2 months).

RESULTS

Characterization of BPV genomes containing linker insertions within the origin. A homologous series of DNA linkers of 5, 10, 15, and 20 bp were inserted into the origin of DNA replication at the *Hpa*I restriction site of the BPV-1 genome such that approximate half-turns and full turns of DNA helix would separate the palindrome. We reasoned that if each half of the palindrome could interact separately with E1 protein, it might be possible to find some biological activity with those insertions that positioned the half-site complexes in the appropriate spatial orientations with respect to each other given the pitch and rise of duplex B-DNA under in vivo conditions. For example, each base insertion at the dyad center rotates the two putative assemblies of protein on the half sites by about 36° and an insertion of 5 bp would rotate the complexes 180° from each other with respect to the wild-type configuration. For a protein complex that itself is in perfect twofold symmetry around this point of rotation, the 180° turn would not change spatial interactions; however, given the potential for an asymmetric assembly, some phasing between half sites might be expected. Thus, 10- and 20-bp linkers may provide a more favorable context to allow the assembly of a competent replication complex than the 5- or 15-bp linker insertions do, as insertions of those lengths restore the phasings between potential contacts on each half site. In contrast, if an E1 protomer must span the dyad axis to bind origin DNA, it would seem unlikely that an insertion would allow much replication activity and, more importantly, a phasing relationship would not be expected. Figure 1 shows the DNA sequences of the synthetic linkers and reveals their sequence relationships. The sequences were designed such that E1 molecules positioned on either side of the separated palindrome would recognize the same flanking linker sequence (5′-CTAG; CTAGC-3′). In addition, the linker sequences have G+C contents close to 50%, as higher percentages may disrupt efficient unwinding of the origin DNA.

The modified BPV genomes harboring the linkers were analyzed for transient replication, transformation efficiency, and stable plasmid copy number in mouse fibroblast C127 cells. In comparison with the wild-type BPV-1 genome, the genomes BPV L.I. 5C, BPV L.I. 10C, BPV L.I. 15C, and BPV L.I. 20C showed levels of transient replication in a down-up-down-up phasing pattern, respectively (Fig. 2). Six days following transfection, the BPV genomes harboring the 10C and 20C linkers transiently replicated to approximately 45% of the wild-type level, whereas the viral genomes containing the 5C and 15C linkers replicated to 6 and 13% wild-type efficiency, respectively. All BPV genomes containing modified origins were able to transform C127 cells, and the extents of transformation paralleled the pattern observed for replication (Table 1). The results shown in Table 1 indicate that insertions of linkers that restore complete turns between the half sites of the dyad axis have higher transformation efficiencies than do those offset by 180°. It appears that the plasmids were not as defective for transformation as they were for replication, especially the BPV L.I. 10C and BPV L.I. 20C genome constructs that transform C127 cells at 80 to 90% of the wild-type level. These transformation data likely reflect in part the need for a certain BPV-1 plasmid copy number for effective levels of transcription of the viral oncogenes. Thus, when amplification of the plasmid is defective, the copy number required for transformation either takes more time to achieve or is too low.

Cell lines were established from the transformed cells by pooling the foci and passaging the cells for approximately 2 months. Low-molecular-weight DNA extracted from these stable cell lines showed that (i) the viral DNA was present as plasmids in all cases and (ii) stable copy number reflected the transient replication capacities (Fig. 3A). That is, BPV L.I. 5C and BPV L.I. 15C showed lower levels of stable plasmid copy number than did BPV L.I. 10C or BPV L.I. 20C. Quantitation of the bands shown in Fig. 3A by PhosphorImager analysis indicated copy numbers to be approximately 18, 90, 44, and 66 for the pools derived from 5C, 10, 15C, and 20C genomes,

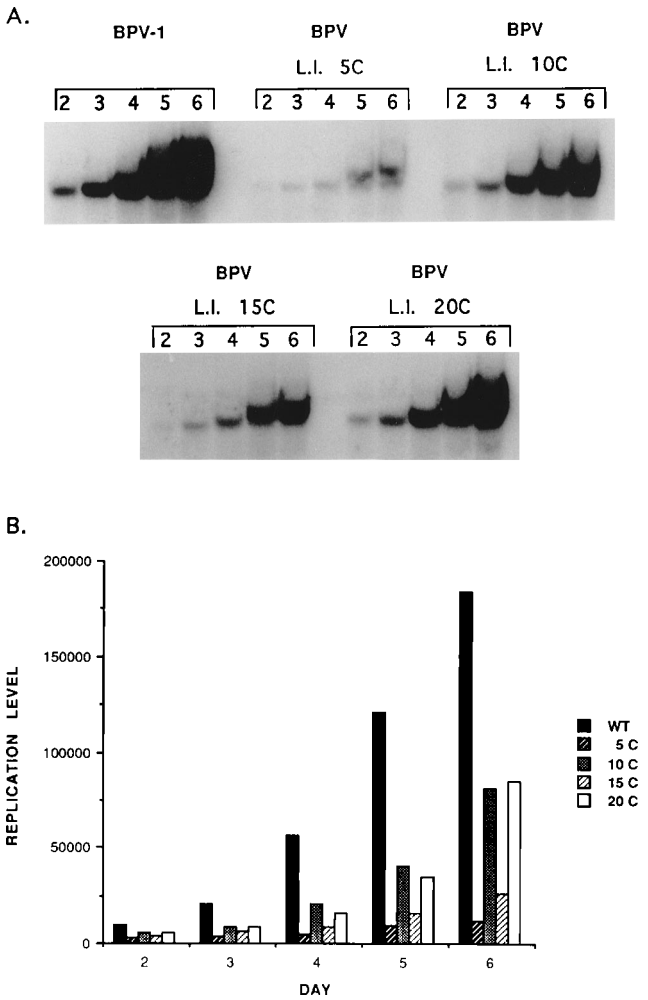

FIG. 2. (A) Transient replication of BPV genomes containing a homologous series of linker insertions within the origin shows the importance of DNA sequence phasing at the palindrome for DNA replication. C127 cells were transfected with 1 μg of viral DNA that was released from the pUC vector by digestion with *Bam*HI. Plasmid DNA was isolated by the method of Hirt and digested with *Eco*RI and *Dpn*I. *Eco*RI linearizes the circular forms of BPV DNA, while *Dpn*I cuts the input DNA, propagated in *E. coli*, that fails to replicate in the animal cells. Samples were subjected to gel electrophoresis, Southern transfer to nitrocellulose, and hybridization with a ^{32}P-labeled nick-translated BPV probe. Samples were taken from days 2 through 6 following electroporation, as indicated above the lanes. (B) Plot of replication levels of BPV genomes harboring linker insertions in comparison with wild-type (WT) replication. Quantitation of the bands detected in the experiment shown in panel A was performed with a PhosphorImager, and the numbers reflect arbitrary readings from the instrument.

respectively, the pooled foci from wild-type BPV-1 DNA being assigned a value of 150 copies. Furthermore, restriction digestion of the extracted DNA verified that the linkers were still present within the origin DNA (Fig. 3B). The 10C, 15C, and 20C linkers contain an *Xba*I restriction site. Therefore, hydrolysis of the plasmid DNA with *Xba*I generates two DNA fragments of 6,132 and 1,813 bp (Fig. 3B). The 5C linker, similar to BPV-1 DNA, does not have the *Xba*I site at the *ori* position; therefore, only the linear form is generated upon *Xba*I digestion.

The phasing pattern for replication displayed by the mutants reflects an intrinsic feature of the origin of DNA replication. It seemed possible that the presence of linker sequences within the viral origin of replication may have influenced the expression of the viral genes dependent upon the nearby viral promoters (P_{7940} and P_{89}), and this in turn may have created the phasing phenomenon. In order to investigate this possibility, the linker sequences were cloned into the context of the pKSO construct. The viral construct pKSO contains 240 bp of origin DNA that includes E2 binding sites 11 and 12, the E1 recognition palindrome, and stretches of A/T-rich sequences. The series of pKSO plasmids were tested for transient replication in the presence of E1 and E2 proteins expressed in *trans* from heterologous promoters (pCG-Eag and pCG-E2). Figure 4 reveals that transient replication with the series of

TABLE 1. Transformation capacities of BPV genomes with linker insertions[a]

BPV plasmid	Foci/μg of DNA in experiment:		
	1	2	3
BPV-1	2,530	2,700	2,540
BPV L.I. 5C	520	680	640
BPV L.I. 10C	2,210	2,340	2,220
BPV L.I. 15C	800	740	710
BPV L.I. 20C	2,070	2,180	2,130

[a] Morphological transformation assays of mouse C127 cells were performed as described in Materials and Methods. The results of three independent experiments are shown.

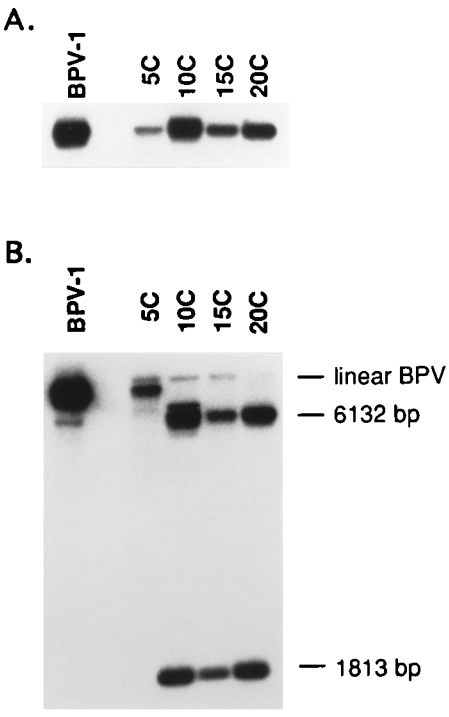

FIG. 3. Detection of plasmid DNA extracted from stably transformed C127 cells. Either wild-type DNA or mutant genomes were used to morphologically transform C127 cells. Pools of foci derived from transfection with each mutant (or the wild type) were made and passaged. Hirt DNA was extracted after serial passage of cells and analyzed by blotting analysis after digestion with restriction enzymes. (A) Detection by autoradiography of a Southern blot of plasmid DNA from DNAs hydrolyzed with *Eco*RI, which linearizes the BPV-1 DNA. (B) Same as panel A except that samples were treated with *Xba*I, which generates two fragments in the 10C, 15C, and 20C genomes because the linker sequence contains an *Xba*I site, as does the BPV-1 genome. For both panels A and B, the equivalent of about 1/20 of the sample from a plate of confluent cells was analyzed for each population of cells.

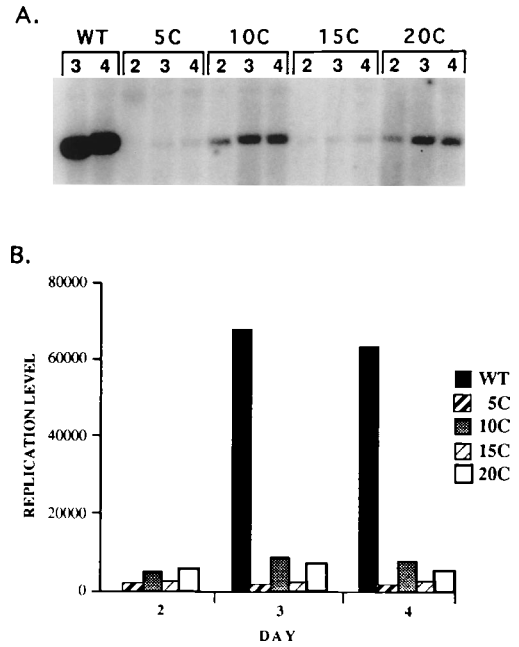

FIG. 4. Importance of DNA spacing between the palindromic half sites at the origin site is detected with replication of minimal origin constructs harboring linker insertions at the *Hpa*I site in transient replication assays. (A) C127 cells were cotransfected with 2 μg of supercoiled replication template (either pKSO L.I. 5C, pKSO L.I. 10C, pKSO L.I. 15C, or pKSO L.I. 20C), 3 μg of E2 expression vector (pCG-E2), and 4 μg of E1 expression vector (pCG-Eag). Plasmid DNA was isolated and processed as described in the legend for Fig. 2. (B) Quantitation of replication levels of the pKSO constructs containing linker insertions and wild-type reporter pKSO. WT, wild type.

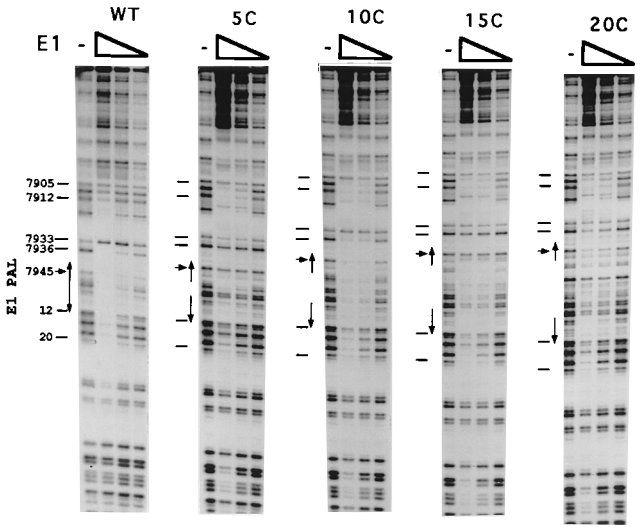

FIG. 5. DNase I footprint analysis of wild-type and mutant genomes bound to E1 protein. The constructs pKSO (WT), pKSO L.I. 5C, pKSO L.I. 10C, pKSO L.I. 15C, and pKSO L.I. 20C were linearized by digestion with *Eco*RI. The DNA ends were labeled with [γ-^{32}P]ATP and digested with *Bam*HI to release the origin fragment (nt 7805 to 100). DNase I protection analysis was performed with the origin templates and purified E1 protein. In each panel, the first lane, marked "−," shows the DNase I pattern of the template DNA alone while the following three lanes represent samples incubated with 120, 80, and 40 ng of E1 protein, respectively. The E1 palindrome (PAL) is indicated by the double-headed vertical arrow next to the pattern for the wild-type origin. The separation of the E1 palindrome by the linker insertions is indicated by the two smaller vertical arrows in subsequent patterns generated by E1 protein bound to substrates 5C, 10C, 15C, and 20C. Numbers next to the wild-type DNase T pattern correspond to BPV-1 genome positions, and the horizontal lines track the DNase-sensitive bands in the corresponding positions. E2 binding site 12 is directly adjacent to the bottom of the E1 palindrome.

pKSO plasmids (pKSO L.I. 5C, pKSO L.I. 10C, pKSO L.I. 15C, and pKSO L.I. 20C) manifested a clear-cut phasing pattern of replication. However, in these experiments it would seem that all of the mutants were considerably more defective than they were in the genome experiments (compare Fig. 4B and 2B). Thus, the distance separating the half sites is playing a major role in the mutants, and the phasing function is superimposed upon this effect. We do not understand this highly reproducible result, which was repeated with different preparations of DNA in different independent experiments. The contrasting results may reflect the synergistic effects of other E2 sites, which may counteract the effects of E1 site mutations in the genomes, or the artificially high expression levels of E1 and E2, which amplify the differences between the wild type and mutants for the measurements with the reporter constructs.

Binding of the E1 protein to the origins containing linkers correlates with the phasing pattern of replication. To explain the replication levels of BPV genomes containing linkers, it was considered that the phasing pattern might be due to the ability of the E1 protein to recognize and cooperatively bind better to those origin sequences possessing linkers that separate the E1 half sites by full turns of helical DNA (10 and 20 bp) as opposed to half-turns (5 and 15 bp). This issue was addressed by performing DNase I protection (footprint) analysis with purified E1 and E2 proteins and substrates of ^{32}P-labeled origin fragments (nt 7805 to 100) containing the linkers. The initial footprint assays were performed by incubating the E1 protein alone with the origin templates at 4°C. Figure 5 shows that the E1 protein (120 ng) binds as previously described in our laboratory to the wild-type origin sequence and that protection spans from nt 7912 to 20 along the bottom strand. Since this protection pattern extends beyond the E1 palindromic sequence, it is likely that E1 is able to oligomerize along the DNA. With smaller amounts of E1 protein (80 and 40 ng), the area of protection narrows to nt 7937 to 11, with some DNase sensitivity arising at nt 7936, 7945, 3, and 6. With all amounts of E1 protein, it appears that nt 7933 remains sensitive to DNase I.

In the presence of 120 ng of E1, all of the template origins containing linkers show protection by E1 protein on the downstream half of the E1 palindrome (nt 4 to 12) closest to E2 binding site 12. For this reason, we believe that the downstream half site contains a primary binding site sequence for the E1 protein and may be the sequence to which E1 initially binds. With respect to the upstream half-site sequence (nt 7942 to 3), the E1 protein is able to bind to the 10C and 20C templates but appears to not properly assemble over 5C and 15C. Nucleotides 7936 and 7945 of 5C and 15C are sensitive to DNase, and it is nt 7945 that is situated at the middle of the upstream half of the palindrome. Nucleotide position 7945 is identified by an arrow in the autoradiograms shown in Fig. 5, and the +/− pattern of protection can be readily observed by using this band as an indicator. Two interesting features of these patterns focus on the protections observed over the linker sequences themselves, and over those nucleotides 5′ to the upstream half of the palindrome (e.g., nt 7933 and 7936). For the 15C insertion mutant, there was strong protection over the linker region, weak protection to the distal palindromic sequences, and no protection at all to those nucleotides 5′ to the upstream site. Similarly, for the 5C insertion there is little or no DNase I protection to the indicator bases 7933 and 7936,

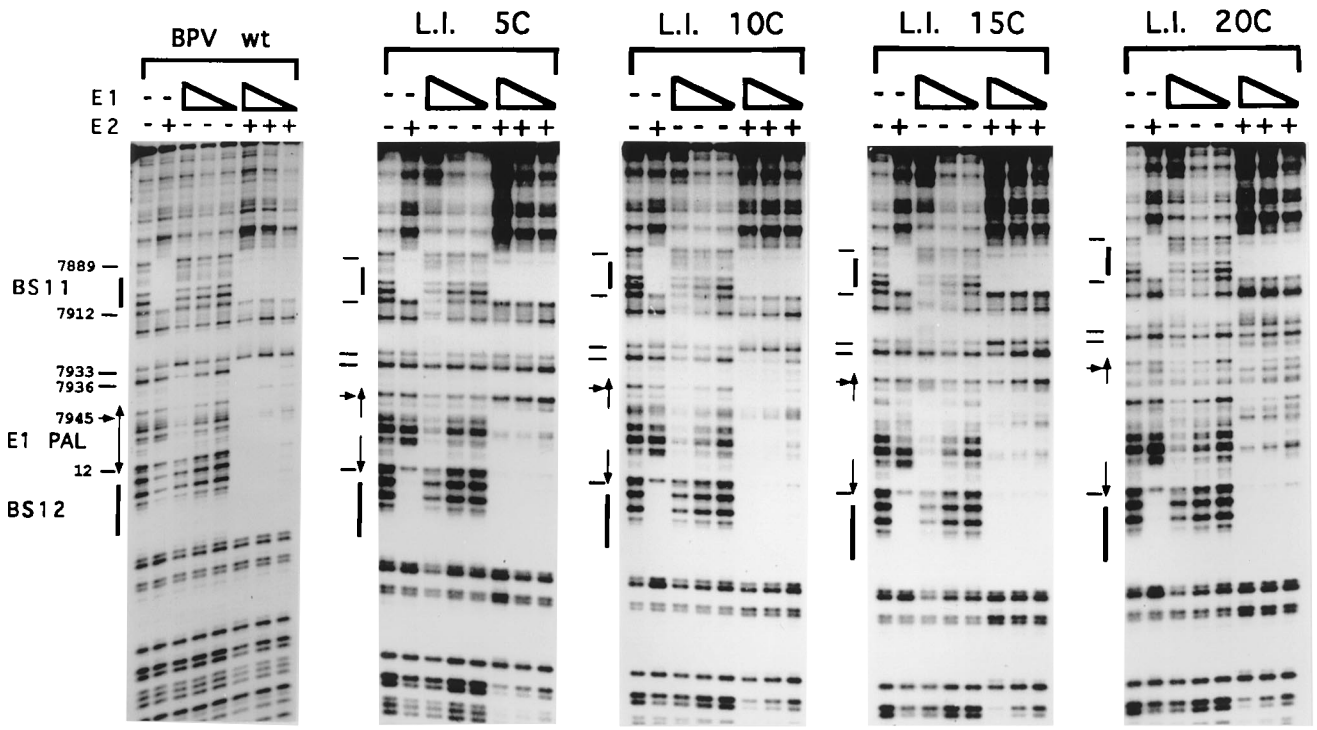

FIG. 6. The E2 protein enhances E1 binding to origins containing linkers. DNase I protection analyses were performed with the DNA templates shown in Fig. 4 and purified E1 and E2 proteins. In all panels, the leftmost lane reveals the DNase I pattern of the template without viral protein. A 50-ng amount of purified E2 protein was utilized either alone or in combination with 120, 80, or 40 ng of purified E1 protein. E2 binding sites (BS) 11 and 12 are indicated by vertical lines, and other markings in the figure are as described in the legend to Fig. 5. In contrast to the conditions used for Fig. 5, the binding assays were performed at 37°C, the temperature at which E2:E1 cooperativity in binding is measured.

while, in contrast, for the 20C insertion, as for the wild-type genome and the 10C insertion, there is some protection over the nucleotides 5′ to the upstream palindrome. It seems likely that the poor binding of E1 to this upstream half of the palindrome is detrimental to the replication of BPV L.I. 5C and BPV L.I. 15C.

The binding of the E1 protein to the 10C template appears to be similar to that to the wild-type origin in that the protection upstream of the origin extends to nt 7905. Therefore, the ability of this construct to achieve a significant amount of replication is predictable. However, the protection pattern conferred by E1 on the 20C insert and its relationship to replication is a bit more problematic. With 120 ng of E1 protein, there appears to be some protection by E1 on the upstream half of the palindrome, and loss of this protection is readily noticeable at 80 and 40 ng of protein. We suspect that an important criterion for the productive protection pattern on this half site is whether DNase sensitivity at nt 7945 is evident. As this base lies in the center of the 5′ site, its protection should be indicative of occupancy. Clearly, at 120 ng of E1 the 20C template does show this protection, while nt 7945 is readily DNase sensitive on the 5C and 15C templates. We conclude that the correlations between E1 binding to the upstream side of the palindrome and replication efficiencies are reasonable. Actual quantitative differences may be better related to higher-order (or different) complexes involving E2 or cellular replication proteins which are subsequently sequestered to this site. In particular, for the 20C construct, which replicates well in the intact genome, other proteins may be needed to stabilize the preinitiation complex.

Figure 6 shows the protection of these synthetic origins in the presence of both E1 and E2. E2 enhanced the binding of E1 to the DNA, although no significant differences in the conclusions drawn from the data shown in Fig. 5 could be made. These binding studies were performed at 37°C, because cooperative interactions between E1 and E2 are optimized at this temperature. Even in the presence of E2, E1 still assembled poorly on the 5C and 15C origins as the upstream half site still showed DNase sensitivity at nt 7945, compared with the wild type. Interestingly, the presence of linker DNA at the palindrome center changes the affinity of E2 for the DNA at binding site 12. We noticed that the E2 footprint over binding site 12 was greater on all of the mutant templates than on the wild type. This is especially noticeable in the absence of E1 (Fig. 6). It has been reported previously that E2 binding to DNA is affected by flanking DNA sequences (20), and we suspect that because E2 bends DNA when it binds (10), the insertions may disrupt (or enhance) conformations in the native DNA which affect binding.

An interesting feature concerning the footprint assays is that, with some of the constructs, the linker sequences were protected by the E1 protein. One might have anticipated that protection by E1 would correspond to the palindrome half sites and that the linker sequence might be sensitive to DNase I, for example, in the 10C construct, where binding is detected on both sides of the palindrome. However, DNase I requires about 10 bp of DNA sequence to interact with DNA, and it cleaves in the minor groove (30). Therefore, it is not surprising that the 10C mutant appears protected over the linker region. Given the above limitations of the DNase I assay, we cannot make any conclusions about protein occupancy in the linker region, particularly for the 10C fragment. However, the 20C linker showed clear DNase I sensitivity over the linker region. One explanation that might explain these patterns is that E1

complexes with E2, E2 occupying binding site 12 and E1 covering the adjacent half site. If this structure is in phase with E1 binding sites on the other side of the dyad axis, further assembly might ensue to occupy the upstream palindromic half site. In the absence of recognition motifs in the 5′ core of the palindrome, E1 might weakly assemble over the linker and flanking regions. The fact that all of the fragments bind E1 on the downstream half-site sequence again suggested that this site might be important for the initial binding of E1.

Nature of the E1 recognition sequence. One objective of modifying the BPV origin DNA and performing DNase I protection studies was to delineate the DNA sequence recognized by the E1 protein. The footprint analyses performed with the origin mutants suggested to us that E1 binds to both halves of the palindrome, with cooperative interactions occurring between proteins assembled on the 5′ and 3′ portions of the dyad. Therefore, it was anticipated that the E1 recognition motif may be contained within the sequence AAPyAAPyAAPy, which is a consensus representation of both sides. We attempted to learn more about the E1 recognition sequences by examining another region of the BPV-1 genome to which the E1 molecule binds.

Previous experiments have shown that the E1 protein bound to sequences within the E2RE1 region (nt 7621 to 7793) between E2 binding sites 8 and 9 (42, 47). Although the early footprint studies were not performed to ascertain a binding site motif, a simple comparison of the sequences at the defined origin of replication and those within E2RE1 revealed a common pentanucleotide element of ATAAT. This pentanucleotide sequence is the largest element common to both regions. The origin half-site sequence AAPyAAPyAAPy does not occur within the E2RE1 region; therefore, the E1 recognition sequence is not likely to be this 9-mer.

There are two ATAAT pentanucleotide sequences within the E2RE1 region at nt 7690 to 7686 and nt 7704 to 7708 and an additional APyAAPy motif at nt 7740 (Fig. 7A). To evaluate whether the sequence ATAAT is necessary for E1 binding, footprint assays were performed with the wild-type E2RE1 sequence and two mutated fragments (P1 and S2), in which the DNA sequences from nt 7684 to 7698 and from nt 7704 to 7718, respectively, were altered by precise deletion substitution, as indicated in Fig. 7A and Materials and Methods. Although these changes alter more than the ATAAT pentanucleotides, the P1 and S2 mutants serve as an initial approximation to assess the involvement of ATAAT in E1 binding. When a ^{32}P-end-labeled DNA fragment from the wild-type E2RE1 region (nt 7632-7821) was used, the E1 protein yielded protection along the top strand at nt 7668 to 7698, 7704 to 7707, and 7712 to 7739 (Fig. 7B). For the P1 mutant sequence, nt 7681 to 7698 were no longer bound by E1 whereas the other regions showed evidence of protection. Conversely, in the mutant template S2, E1 no longer protected nt 7704 to 7708. Therefore, using the logic of consensus sequence matching and the results of these binding experiments, we suggest that the sequence APyAAPy is an element for binding BPV E1. We do not suggest that this motif is the sole determinant for E1 binding, as it is likely that the context of DNA sequences flanking the site may have strong effects upon the DNA:protein complex and other specific features, such as DNA distortions, may affect affinity. For example, such effects have been determined for E2:DNA protein complexes (19). Moreover, E1 interacts nonspecifically with DNA (46), and while the parameters of specific to nonspecific binding have not been carefully measured, it would appear that, at high concentrations of E1, assembly of artificially large complexes can be coordinated by weak sites or by the origin sequences themselves (45).

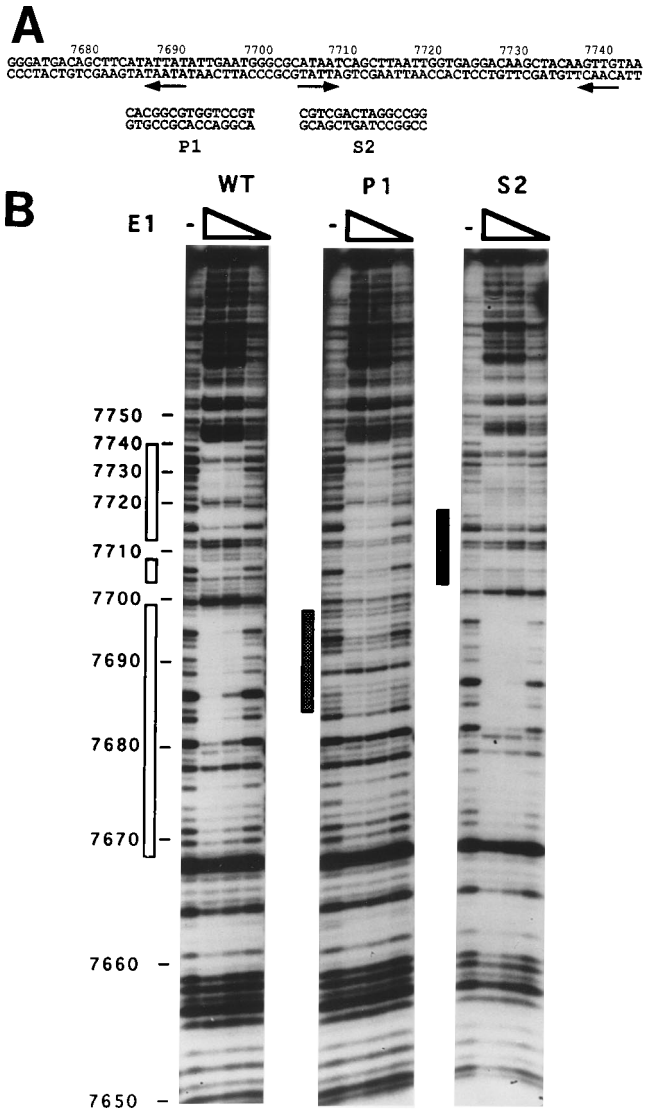

FIG. 7. (A) Diagram of E2RE1 region where the E1 protein binds. The arrows mark the ATAAT pentanucleotide sequences and a closely related ACAAC sequence. The mutations utilized to replace the ATAAT element are designated P1 and S2, and the replacement sequences are indicated. Numbers indicate nucleotide positions. (B) DNase I protection analysis reveals the binding of the E1 protein to sequences within the E2RE1 region. In each panel, the first lane represents the DNase pattern of the template DNA. The following three lanes show the DNase I patterns after binding of 120, 80, and 40 ng of E1 protein, respectively. The numbers at the left indicate BPV-1 nucleotide positions. The open boxes to the left of the wild-type (WT) footprint demark the extent of protection, while the shaded boxes to the left of the P1 and S2 patterns mark the positions of the mutated DNA.

DISCUSSION

A series of synthetic linkers were inserted at the unique *Hpa*I restriction site (nt 3) of the BPV-1 genome. The ability of these modified genomes to replicate implied that the E1 recognition palindrome (nt 7942 to 12) could be separated into functional half sites that are each competent for binding the E1 protein. Although replication of most of the mutant genomes was compromised, none of the insertions were able to completely abolish replication of the viral genome in mouse fibroblast C127 cells.

These observations differ from two earlier reports which

showed that the insertion of 8-bp *Xho*I or *Sal*I linkers at the *Hpa*I restriction site (nt 3) completely abolishes replication of the BPV genome (34, 39). One possibility might be that the sequence of the linker influences the outcome in some way. For example, we have found that two 8-bp linkers (*Xho*I, CCTCGAGG; *Spe*I, GACTAGTC), when inserted into full-length genomes, differ from each other by about fourfold in their replication phenotypes (data not shown). Alternatively, although this and the latter idea are not mutually exclusive, the failure to observe replication of those modified viral genomes tested in one study (39) may be attributable to the manner in which the experiments were performed. Instead of transfecting C127 cells with only the modified viral genome and allowing that genome to rely upon its own production of E1 and E2 for replication, Ustav and Stenlund (39) performed a cotransfection with the modified genome and a BPV genome still attached to the backbone vector (pMLBPV) to provide viral proteins in *trans*. A potential reason for the difference from our results with this assay is that it is rather difficult to keep two different plasmids replicating in C127 cells, especially when the two genomes replicate at different efficiencies and competition for E1 and E2 is manifest (our unpublished observations). A selection is likely to occur such that one of the two genomes will be lost. The lost genome is likely to be the one containing the linker because of its poor ability to replicate. Furthermore, we initially reported that, with the cell-free replication system, a 14-bp *Nco*I linker insertion completely destroyed in vitro replication (45). With improved specific activity of E1 and E2, it has been observed that such mutants do retain 10 to 15% of wild-type activity; moreover, the phasing relationships that we obtained for the genome templates described in this report were also detected in a cell-free system (45a). Consistent with the results reported here, Le Moal et al. (16) have recently shown that a BPV-1 replication reporter which contained a 10-bp *Bam*HI linker inserted at the *Hpa*I site replicated at 10 to 20% the efficiency of the wild type.

Although the linker insertions did not destroy replication, they caused a readily detectable reduction in replication proficiency that confirms the importance of the E1 palindrome as a *cis* element necessary for optimal replication. In addition, the robust replication of the reporter constructs and genomes with 10- and 20-bp linker insertions suggests that more than one E1 molecule binds to the origin palindrome sequence. It is not likely that a single protein molecule can extend across 10 or even 20 bp of linker DNA, binding to a split recognition motif, and continue to function in a significant manner. The phasing pattern of replication observed among mutants harboring the linker insertions implies that the E1 monomers (or multimers) do not independently assemble and function on each half of the separated palindrome. If there had been independent assembly, then perhaps all the modified genomes would have displayed similar levels of replication. It is our conclusion from these results that for efficient DNA replication, E1 molecules on one half site of the palindrome must be able to interact with molecules on the other half site at some point in the initiation process.

The DNase I protection (footprint) experiments support the above ideas. BPV origins containing the 10- and 20-bp linkers showed protection by E1 over both separated E1 half sites. In contrast, BPV origins with the 5- and 15-bp linker insertions replicated poorly and did not show binding of E1 to the upstream half site in a manner similar to binding with the wild type. This is again consistent with the notion that separation of the E1 palindrome by ½ and 1 turns of DNA disrupts the cooperative binding of E1 and results in an inefficient replication complex.

The footprint studies showed that the E2 protein enhances the binding of E1 to the origins containing linkers. In the presence or absence of E2, however, all of the modified origins bound E1 to the palindrome half site adjacent to E2 binding site 12. This phenomenon implies that a binding site for the E1 protein is contained within this half site and that E1 assembly toward a higher-order prereplication complex may begin specifically within this half of the palindrome. It seems logical to consider that the E2 protein initially orients and stabilizes E1 at the adjacent half site and then additional E1 molecules further oligomerize to create a competent replication assembly. A similar conclusion was inferred by Gillette et al. (8), who found that in the presence of E2, protection was first detected over the 3′ half of the E1 binding site dyad axis at low E1 concentrations and then detected on the 5′ sequences at higher concentrations.

The apparent phasing effects created by distance between the half sites detected by our assays raise some interesting questions with regard to how E1 monomers might assemble on the origin site. First, we wonder how the two different arrays (on the 5′ and 3′ halves) might communicate with each other. A 10-bp insertion separates the half sites by about 34 Å (3.4 nm); in scales relevant for proteins, this is a large distance, and it is perhaps twice this length for the 20-bp insertion. A length spanning 10 or 20 bp is perhaps too small for a DNA loop, and a deformation of the B-DNA that would allow for a melting of such distances would ruin the modular 10-bp relationships found. Similar close-range phasing relationships have also been detected by transcription factors which assemble upon the beta interferon promoter (36a). For the beta interferon gene, the PRD IV element of the promoter contains abutting ATF2, Jun, and HMGI binding sites. Insertions of 5 bp between the ATF2 and c-*jun* site inactivate the element, but a 10-bp insertion restores activity. In that situation also, no specific resolution to the question of how the short linker sequences are arranged has been uncovered. In their classical study focused on λ repressor proteins bound to distal sites on DNA and the effects of distance between binding sites on DNA and cooperative binding, Hochschild and Ptashne (12) varied λ repressor binding sites between 4.6 and 6.4 turns of DNA. In that study of the λ repressor, it was calculated that the energetic cost of creating such a bend would be small but would dramatically increase with short DNA loops.

In order for two proteins to bind cooperatively and be separated by only 10 or 20 bp, perhaps a flexible "accordion" domain might bring the monomers in contact, as has been proposed for the carboxy-terminal domain of the alpha subunit of *Escherichia coli* RNA polymerase in its orientation- and distance-dependent interactions with the CAP protein (1). Such a flexible domain might be restricted to protein-protein interactions presented on the same face of the helix and may not be capable of wrapping around the duplex. For a DNA helicase which might need to multimerize and form two shells around the DNA, one on each side of a palindromic sequence, each monomer might have such a flexible domain which could stretch across the linker space to its appropriate partner. Another idea, which is perhaps suggested by the apparent DNase I protection conferred by E1 on linker sequences in the mutant DNAs examined here, is that E1 actually assembles on the linkers, after which, if the 5′ sequences (upstream palindromic DNA) are in register with this E1:linker DNA complex, further specific assembly takes place over the distal site. We are hesitant, however, to conclude that E1 actually does bind to the linker DNA, given the limitations of the DNase I assay as discussed in Results, and find the flexible-protein notion or DNA bend hypothesis simpler, only because it is hard to con-

ceive of two specific complexes communicating through the proteins coordinated over the linker region. However, we point out that in some other papillomaviruses clear palindromic sequences do not occur at the origin sites, and it may be that in those situations one half site coordinates assembly of the other E1 molecules over relatively weak E1 DNA contact sites. For the human papillomavirus type 11 replication origins (31), we have found that mutation of the 3′ side of the E1 palindrome has a greater effect upon replication than does mutation of the 5′ portion, consistent with this notion. For linker insertions, perhaps the 5′ palindromic DNA must be in register with the 3′ sequences so as to not compete nonproductively with assembly over the linker region and 3′ sites.

A related problem concerns the mechanism by which the inserted linker sequences can be melted for initiation, if the linker region is not occupied by a functional helicase assembly. It is difficult to speculate how such an insertion would be dealt with, as essentially nothing in detail is understood about how helicases function. Recently presented image reconstructions of the RuvB helicase show that duplex DNA, bound to this particular helicase assembly, runs through the center of a doughnut-shaped structure (35). Therefore, left-handed twisting of the right-handed DNA helix may unwind the molecule on both sides of the protein once helicase activity begins.

Another goal of our studies was to provide some insights into what DNA recognition sequences interact with the E1 DNA binding domain. Previous results from our laboratory had shown that E1 can bind to BPV-1 DNA sequences located between nt 7621 and 7793 (42, 47). Inspection of these sequences revealed that a repeating pentanucleotide sequence of ATAAT (or APyAAPy) is present in this region and is also found within the origin dyad symmetry region. We have mutated the upstream binding sites by a substitution mutation, and the results of DNase I protection analysis show that these regions do determine E1 DNA affinity. The mutations introduced are three times larger than the putative recognition motif that they encompass, and more mutagenesis coupled with DNA binding will be necessary to determine precisely the important recognition nucleotides. However, we believe it likely that some version of this consensus repeating sequence is involved in E1 DNA binding. The P1 mutation, in which one of the APyAAPy repeats is mutated, has the most profound effect upon decreasing protein-DNA interactions throughout the region. In contrast, mutation over the S2 region, which contains another pentanucleotide, left protection over the nt 7690 region intact while destroying protection locally. These differences imply that either DNA context or other nucleotide recognition specificities are required for strong binding.

Figure 8 displays some possible ways in which the pentanucleotide repeats may aid in the assembly of E1 over the replication origin dyad axis. In line I of Fig. 8, we show all of the pentamers, and in this arrangement there is a 2-bp overlap with all of the recognition motifs. This type of arrangement would require a wrapping of the protomers around the DNA, and the 5′ side of the dyad axis would probably be a larger structure than what would be present on the 3′ side, a feature perhaps indicated by the DNase I protection patterns observed (8, 40, 45). In line II of Fig. 8, we show a symmetrical compact structure of repeats more reminiscent of the pentanucleotide organization known to be important for SV40 and polyomavirus origin function. In line III of Fig. 8, the repeats are drawn without overlap and, as shown in line I, the structural organization around the dyad axis may be predicted to be asymmetrical. It is, of course, possible that all of these assemblies can form and that some are abortive with respect to initiation while others could function. Furthermore, we suspect that E2 affects

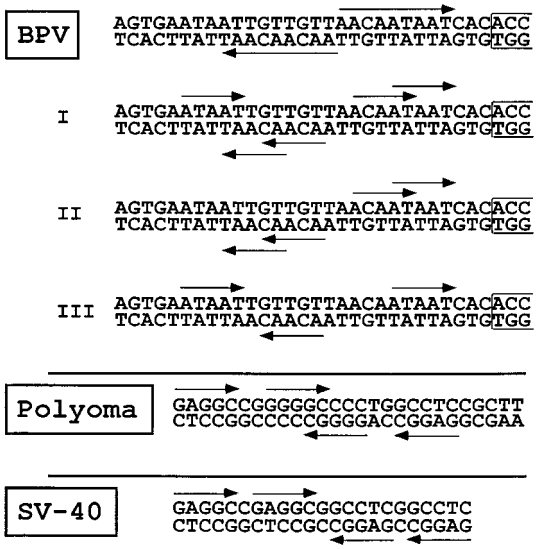

FIG. 8. Diagrammed at the top are three possible ways that the E1 protein may bind to an APyAAPy sequence within the BPV origin of replication. Line I shows all of the pentamer motifs, line II shows the motifs in a symmetrical pattern organized with twofold symmetry, and line III shows the motifs in a nonoverlapping array. At the bottom, the GPuGGC repeats in polyomavirus are shown in contrast to the SV40 GAGGC repeats. These repeats define the T-antigen binding sites in these viruses.

the pathway of assembly, as a strong DNase I-hypersensitive site appears in the middle of the E1 dyad axis for the patterns found in the top strand only when E2 is present (references 45 and 47 and our unpublished observations). For comparison with these models, we show in Fig. 8 the relative arrangements of the pentanucleotide repeats of the sequence GPuGGC found at the SV40 and polyomavirus origins of replication. With the polyomavirus, where large T antigen binds weakly, it is not clear if all of the pentamers are occupied at once (4). However, for the SV40 origin, productive assembly of T antigen requires binding to each of the GAGGC repeats, two of which on each side of the dyad axis are believed to coordinate hexamers (2). Moreover, the spacing between the individual pentanucleotide repeats on each side of the palindrome is known to be critical for SV40 (29), although it has not been determined whether a phasing relationship between the right half and left half of the palindrome is operative in this system, as appears to be the case for BPV-1. What seems to be apparent from inspection of the sequence of the BPV-1 origin is that the pentanucleotide repeats cannot be arranged in a manner identical to the arrangement of GPuGGC repeats of SV40 or polyomavirus. Line II in Fig. 8 is most reminiscent of (or homologous to) the DNA organization of the other papovaviruses, and further work is necessary to learn just how similar the initiation complexes on the respective DNAs are.

ACKNOWLEDGMENTS

This work was supported by funds from the NIH through grants CA-42414, CA-30490, and ES-01896.

REFERENCES

1. **Blatter, E., W. Ross, H. Tang, R. L. Gourse, and R. H. Ebright.** 1994. Domain organization of RNA pol α subunit: C-terminal 85 amino acids constitutes a domain capable of dimerization and DNA binding. Cell **78:** 889–896.
2. **Borowiec, J. A., F. B. Dean, P. A. Bullock, and J. Hurwitz.** 1990. Binding and unwinding—how T antigen engages the SV40 origin of DNA replication. Cell **60:**181–184.

3. **Clertant, P., and I. Seif.** 1984. A common function for polyoma virus large-T and papillomavirus E1 proteins? Nature (London) **311:**276–279.
4. **Cowie, A., and R. Kamen.** 1986. Guanine nucleotide contacts within viral DNA sequences bound by polyomavirus large T antigen. J. Virol. **57:**505–514.
5. **Deb, S., S. Tsui, A. Koff, A. L. DeLucia, R. Parsons, and P. Tegtmeyer.** 1987. The T-antigen-binding domain of the simian virus 40 core origin of replication. J. Virol. **61:**2143–2149.
6. **DeLucia, A. L., B. A. Lewton, R. Tjian, and P. Tegtmeyer.** 1983. Topography of simian virus 40 A protein-DNA complexes: arrangement of pentanucleotide interaction sites at the origin of replication. J. Virol. **46:**143–150.
7. **Dvoretzky, I., R. Shober, S. K. Chattopadhyay, and D. Lowy.** 1980. A quantitative in vitro focus assay for bovine papilloma virus. Virology **103:**369–375.

7a. **Fouts, E., and M. Botchan.** Unpublished data.

8. **Gillette, T. G., M. Lusky, and J. A. Borowiec.** 1994. Induction of structural changes in the bovine papilloma virus type 1 origin of replication by the viral E1 and E2 proteins. Proc. Natl. Acad. Sci. USA **91:**8846–8850.
9. **Hawley-Nelson, P., E. Androphy, D. R. Lowy, and E. J. Schiller.** 1988. The specific DNA recognition sequence of the bovine papillomavirus E2 protein is an E2-dependent enhancer. EMBO J. **7:**525–531.
10. **Hegde, R. S., S. R. Grossman, L. A. Laimins, and P. B. Sigler.** 1992. Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. Nature (London) **359:**505–512.
11. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
12. **Hochschild, M., and M. Ptashne.** 1986. Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. Cell **44:**681–687.
13. **Holt, S. E., G. Schuller, and V. G. Wilson.** 1994. DNA binding specificity of the bovine papillomavirus E1 protein is determined by sequences contained within an 18-base-pair inverted repeat element at the origin of replication. J. Virol. **68:**1092–1102.
14. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell **42:**559–572.
15. **Kuo, S. R., J. S. Liu, T. R. Broker, and L. T. Chow.** 1994. Cell-free replication of the human papillomavirus DNA with homologous viral E1 and E2 proteins and human cell extracts. J. Biol. Chem. **269:**24058–24065.
16. **Le Moal, M. A., M. Yaniv, and F. Thierry.** 1994. The bovine papillomavirus type 1 (BPV1) replication protein E1 modulates transcriptional activation by interaction with BPV1 E2. J. Virol. **68:**1085–1093.
17. **Li, R., and M. R. Botchan.** 1993. The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate in vitro BPV-1 DNA replication. Cell **73:**1207–1221.
18. **Li, R., and M. R. Botchan.** 1994. Acidic transcription factors alleviate nucleosome-mediated repression of BPV-1 DNA replication. Proc. Natl. Acad. Sci. USA **91:**7051–7055.
19. **Li, R., J. Knight, G. Bream, A. Stenlund, and M. Botchan.** 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. Genes Dev. **3:**510–526.
20. **Li, R., L. Yang, E. Fouts, and M. R. Botchan.** 1993. Site specific DNA binding proteins important for replication and transcription have multiple activities. Cold Spring Harbor Symp. Quant. Biol. **58:**403–413.
21. **Lusky, M., J. Hurwitz, and Y.-S. Seo.** 1994. The bovine papillomavirus E2 protein modulates the assembly but is not stably maintained in a replication-competent multimeric E1-replication origin complex. Proc. Natl. Acad. Sci. USA **91:**8895–8899.
22. **MacPherson, P., L. Thorner, and M. R. Botchan.** 1994. The bovine papilloma virus E1 protein has ATPase activity essential to viral DNA replication and efficient transformation in cells. Virology **204:**403–408.
23. **Mastrangelo, I., P. Hough, J. Wall, M. Dodson, F. Dean, and J. Hurwitz.** 1989. ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. Nature (London) **338:**658–662.
24. **McBride, A. A., J. C. Byrne, and P. M. Howley.** 1989. E2 polypeptides encoded by bovine papillomavirus 1 form dimers through the carboxy-terminal DNA-binding domain: transactivation is mediated through the amino-terminal domain. Proc. Natl. Acad. Sci. USA **86:**510–514.
25. **Mohr, I. J., R. Clark, S. Sun, E. J. Androphy, P. MacPherson, and M. R. Botchan.** 1990. Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator. Science **250:**1694–1699.
26. **Muller, F., Y.-S. Seo, and J. Hurwitz.** 1994. Replication of bovine papillomavirus type 1 origin-containing DNA in crude extracts and with purified proteins. J. Biol. Chem. **269:**17086–17094.
27. **Park, P., W. Copeland, L. Yan, T. Wang, M. Botchan, and I. J. Mohr.** 1994. The cellular DNA polymerase α-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase. Proc. Natl. Acad. Sci. USA **91:**8700–8704.
28. **Parsons, R. E., J. Stenger, S. Ray, R. Welker, M. Anderson, and P. Tegtmeyer.** 1991. Cooperative assembly of simian virus 40 T-antigen hexamers on functional halves of the replication origin. J. Virol. **65:**2798–2806.
29. **Parsons, R. E., and P. Tegtmeyer.** 1992. Spacing is crucial for coordination of domain function within the simian virus 40 core origin of replication. J. Virol. **66:**1933–1942.
30. **Rhodes, D., and A. Klug.** 1980. Helical periodicity of DNA determined by enzyme digestion. Nature (London) **286:**573–578.
31. **Russell, J. A., and M. R. Botchan.** 1995. *cis*-acting components of human papillomavirus (HPV) DNA replication: linker substitution analysis of the HPV type 11 origin. J. Virol. **69:**651–660.
32. **Schvartzman, J. B., S. Adolph, L. Martin-Parras, and C. L. Schildkraut.** 1990. Evidence that replication initiates at only some of the potential origins in each oligomeric form of bovine papillomavirus type 1 DNA. Mol. Cell. Biol. **10:**3078–3086.
33. **Seo, Y.-S., F. Muller, M. Lusky, and J. Hurwitz.** 1993. Bovine papilloma virus (BPV)-encoded E1 protein contains multiple activities required for BPV DNA replication. Proc. Natl. Acad. Sci. USA **90:**702–706.
34. **Spalholz, B. A., A. A. McBride, T. Sarafi, and J. Quintero.** 1993. Binding of bovine papillomavirus E1 to the origin is not sufficient for DNA replication. Virology **193:**201–212.
35. **Stasiak, A., I. R. Tsaneva, S. C. West, C. J. B. Benson, X. Yu, and H. Egelman.** 1994. The *Escherichia coli* RuvB branch migration protein forms double hexameric rings around DNA. Proc. Natl. Acad. Sci. USA **91:**7618–7622.

35a. **Stenlund, A.** Personal communication.

36. **Tanaka, M., and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. Cell **60:**375–386.

36a. **Thanos, D., and T. Maniatis.** Personal communication.

37. **Thorner, L., N. Bucay, J. Choe, and M. Botchan.** 1988. The product of the bovine papillomavirus type 1 modulator gene (M) is a phosphoprotein. J. Virol. **62:**2474–2482.
38. **Thorner, L., D. Lim, and M. R. Botchan.** 1993. DNA-binding domain of bovine papillomavirus type 1 E1 helicase: structural and functional aspects. J. Virol. **67:**6000–6014.
39. **Ustav, M., and A. Stenlund.** 1991. Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. EMBO J. **10:**449–457.
40. **Ustav, M., E. Ustav, P. Szymanski, and A. Stenlund.** 1991. Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. EMBO J. **10:**4321–4329.
41. **Wessel, R., J. Schweizer, and H. Stahl.** 1992. Simian virus 40 T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. J. Virol. **66:**804–815.
42. **Wilson, V. G., and J. Ludes-Meyers.** 1991. A bovine papillomavirus E1-related protein binds specifically to bovine papillomavirus DNA. J. Virol. **65:**5314–5322.
43. **Yang, L., and M. Botchan.** 1990. Replication of bovine papillomavirus type 1 DNA initiates within an E2-responsive enhancer element. J. Virol. **64:**5903–5911.
44. **Yang, L., C. B. Jessee, K. Lau, H. Zhang, and L. F. Liu.** 1989. Template supercoiling during ATP-dependent DNA helix tracking: studies with simian virus 40 large tumor antigen. Proc. Natl. Acad. Sci. USA **86:**6121–6125.
45. **Yang, L., R. Li, I. J. Mohr, R. Clark, and M. R. Botchan.** 1991. Activation of BPV-1 replication *in vitro* by the transcription factor E2. Nature (London) **353:**628–633.

45a. **Yang, L., R. Mendoza, and M. Botchan.** Unpublished data.

46. **Yang, L., I. Mohr, E. Fouts, D. A. Lim, M. Nohaile, and M. Botchan.** 1993. The E1 protein of the papillomavirus BPV-1 is an ATP dependent DNA helicase. Proc. Natl. Acad. Sci. USA **90:**5086–5090.
47. **Yang, L., I. Mohr, R. Li, T. Nottoli, S. Sun, and M. Botchan.** 1991. The transcription factor E2 regulates BPV-1 DNA replication *in vitro* by direct protein-protein interaction. Cold Spring Harbor Symp. Quant. Biol. **56:**335–346.