LIU YANG AND MICHAEL BOTCHAN\*

Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, 401 Barker Hall, University of California, Berkeley, California 94720

Received 29 June 1990/Accepted 24 August 1990

When bovine papillomavirus transforms cells in vitro, it maintains its genome as a multicopy nuclear plasmid. Plasmid DNA extracted from such transformed cells was analyzed by the two-dimensional gel electrophoresis technique of Brewer and Fangman (B. Brewer and W. Fangman, Cell 51:463-471, 1987). The replication intermediates detected in these assays were found to be the sums of the oligomeric and monomeric forms of the replicating plasmids. The multimeric DNAs were shown by field inversion gel electrophoresis and partial restriction digestion to be head-to-tail concatemers of the monomeric forms. Furthermore, the multimers progressed in size by steps of one monomer, indicating that they did not arise by replication segregation mistakes of the unit length, which would predict a ladder spaced by integrals of two monomers. To map the plasmid DNA replication origin, the replication intermediates of the monomers were isolated by successive sucrose gradient centrifugation and then examined by the two-dimensional gel electrophoresis method. The patterns detected show that bovine papillomavirus type 1 replicates in these cells bidirectionally and that one replication origin site in the viral genome is utilized. By employing several restriction enzymes and specific viral DNA probes to dissect the replication intermediates, we were able to map the origin of initiation site with some precision. The initiation site, which maps to bovine papillomavirus type 1 DNA position 7730  $\pm$ 100 bp, places the origin within that region of the viral upstream regulatory region which contains the major cluster of transcription factor E2-binding sites, E2RE1. Thus, the actual viral plasmid origin of replication maps near, but outside, genetic elements previously shown to be important for plasmid maintenance.

The chromosomes of eucaryotic cells initiate DNA replication at multiple sites termed origins of replication. Numerous studies of the initiation patterns of eucaryotic DNA replication have demonstrated that these initiation sites are regulated both spatially and temporally. Moreover, the spatial pattern of initiation varies in dividing cells within a single species (see references 15 and 42 for reviews). In contrast to our detailed knowledge about metazoan transcriptional promoter-enhancer structures in DNA, very little is known about just how much genetic information is required to specify a replication origin in the chromosomes. Some evidence supports the notion that these sequences may not be particularly "hard-wired," as any DNA fragment can replicate after injection into *Xenopus* eggs (24). However, genetic studies of DNA replication in yeast cells are consistent with trans-acting factors binding to specific sites contained within the autonomous replication sequence elements (20). Moreover, these autonomous replication sequence elements function as replication origins in vivo (4, 19). It is interesting that this conserved A+T-rich core region in the autonomous replication sequence elements is not known to be the recognition site of a DNA-binding protein, though such recognition could clearly be provided for by auxiliary flanking sequences.

DNA replication studies with simian virus 40 (SV40) and polyomavirus provide the most compelling arguments that eucaryotic origin DNA sequences are the targets for sitespecific recognition by initiator proteins. The SV40-encoded large T antigen binds to a specific origin sequence and starts the assembly of the initiation complex in concert with a variety of cellular factors. The helicase activity of T antigen

Viruses that establish themselves in a latent or quasi-latent state within a cell provide interesting models for the study of the regulation of DNA replication in animal cells. Epstein-Barr virus (EBV) establishes itself in transformed human lymphoblasts as a multicopy nuclear plasmid. In these cells, the origin of plasmid replication is a bipartite structure consisting of a series of 20 30-bp repeats and a dyad symmetry element which contains four partial copies of the tandem repeats. These repeats are spaced about 1 kb apart and are both required for stable plasmid maintenance in recombinant DNA vectors derived from EBV. The repeats are the sites for the binding of the EBV-encoded EBNA-1 protein, which is the only EBV protein required in trans for EBV ori replication (33). It has recently been shown that EBV DNA replication initiates in this region (37). The bovine papillomavirus type 1 (BPV-1) also establishes itself in transformed cells as a stable nuclear plasmid. In both systems the plasmid DNA doubles in each round of cellular DNA replication (in most of the cells of a population in

unwinds the double-stranded DNA at this region; this unwinding is then followed by nascent DNA strand synthesis (see references 8 and 40 for reviews). The basic mechanisms of replication initiation and elongation appear to have been remarkably conserved from procaryotes to these eucaryotic viruses. Thus, for all of the most thoroughly studied replicons, initiation of DNA synthesis is preceded by assembly of a protein-DNA complex at a specific origin sequence. The origin region is then unwound by a DNA helicase (5). While SV40 and polyomavirus replicons have been useful models for studying the initiation and elongation phases of DNA replication, these viruses undergo multiple rounds of DNA synthesis in a given cell cycle. Thus, these systems may not be suited for examining the controls which govern once-percell-cycle DNA replication.

<sup>\*</sup> Corresponding author.

vitro), and thus mechanisms must exist for controlling the replication of the plasmid in coordination with the cell cycle.

BPV-1 has a double-stranded circular DNA with a genome size of about 8 kb. The virus causes cutaneous fibropapillomas in cattle and fibroblastic tumors in horses. The creation of a benign tumor is a natural part of the virus's life cycle, and in these transformed cells the viral DNA maintains itself as a nuclear plasmid (23). Although no cell culture system is available to examine the process of vegetative replication, the availability of cell lines in which BPV-1 can be maintained in a stable plasmid form (25) has facilitated the investigation of replication in the latent form of the virus. BPV-1 can transform fibroblasts from a variety of species, including mice, hamsters, and cattle (22). Typically, the viral plasmids maintain themselves at a copy number of 100 to 300 per cell. The E-1 open reading frame of BPV-1 encodes a 68-kDa nuclear protein (40a), and prior genetic studies based on both transient and stable replication assays implicate this protein as one of the important viral proteins required for plasmid replication (3, 30). As this protein shares significant structural (9) and functional (unpublished observations) homologies to the SV40 large T antigen, it is attractive to hypothesize that the E1 protein plays a direct role in BPV-1 replication. The replication activity of E1 may be regulated by a gene called M which encodes for a truncated form of the E1 protein (3, 30, 41; Sun et al., submitted). Two sequence elements, named plasmid maintenance sequence 1 (PMS-1) and PMS-2, have been identified by using recombinant vectors containing DNA fragments of BPV-1 linked to a selectable neomycin resistance gene. Such vectors, upon selection, maintain themselves as supernumerary plasmids in BPV-1-transformed cells but integrate into chromosomes in uninfected cells. These experiments map elements which contain sufficient portions of the sequence for plasmid maintenance and transient amplification (29, 43) in artificial constructs. Nevertheless, they do not necessarily include or contain those sequences actually used as a replication origin in the intact viral plasmid. In the R6K plasmid of Escherichia coli, for example, the minimal replication origin required for stable maintenance consists of the 400-bp  $\gamma$  origin (21). However, in the intact plasmid, three plasmid origins ( $\alpha$ ,  $\beta$ , and  $\gamma$ ) have been mapped, and  $\gamma$  is relatively inactive (11). The  $\gamma$  origin is apparently an enhancer which initially binds the replication initiator protein and subsequently transfers the activity to weak sites within  $\beta$  or  $\alpha$  (35). In the absence of  $\alpha$  or  $\beta$ ,  $\gamma$  is apparently sufficient, but neither  $\alpha$  nor  $\beta$  is sufficient without  $\gamma$ .

The direct visualization of BPV-1 replication intermediates by electron microscopy is difficult, as the plasmid DNA represents only a minute fraction of the total cellular DNA. However, the electron microscopy analysis of Waldeck et al. (43) supports the notion that replication initiation occurs uniquely within or near the defined PMS-1 region in BPV-1. However, that mapping was based upon an alignment of eye forms within two sets of 10 replicating molecules. The recently developed two-dimensional (2D) gel electrophoresis method of Brewer and Fangman (4) can in practice be used to accurately define the point of replication initiation in BPV-1. We report here that in ID13 cells (a transformed mouse cell line) only one origin of BPV-1 DNA replication is detected. Surprisingly, the initiation site as mapped by this technique places the site neither at PMS-1 nor at PMS-2 but rather at a site about 300 to 400 bp downstream of the 3' border of the PMS-1 element. Spalholz et al. (38) had originally defined a region of BPV-1 within the viral upstream regulatory region (URR) as an E2-responsive enhancer fragment (E2RE1). Thus, BPV-1 sequences from positions 7611 to 7808 contain an enhancer element which can stimulate transcription of the heterologous SV40 promoter in an orientation- and position-independent manner in the presence of the BPV-1-encoded E2 protein. This E2RE1 region contains four E2-binding sites: sites 7, 8, 9, and 10 (see Fig. 8B and references 2, 16, 27, and 34). Replication initiation maps to BPV-1 position 7730  $\pm$  100 bp, thus placing the origin within this regulatory region.

# MATERIALS AND METHODS

**Cells.** ID13 cells (mouse C127 cells transformed with BPV-1) were obtained from P. Howley. VI216 cells were described previously (41), and WHZ-10 cells were obtained from S. Burnett and U. Pettersson (Uppsala, Sweden). The cells were maintained in Dulbecco modified Eagle medium (GIBCO) supplemented with 5% fetal bovine serum and 5% newborn calf serum.

**Preparation of DNA and sucrose gradient centrifugation.** BPV-1 DNA was isolated from subconfluent ID13 cells by the Hirt method (18). The DNA was then extracted once with phenol-chloroform and once with chloroform and was precipitated with ethanol. To separate multimeric forms of BPV-1, the DNA was loaded on a sucrose gradient (10 to 30%) in 1 M NaCl-20 mM Tris (pH 7.7)-1 mM EDTA and centrifuged in a Beckman SW41 rotor at 36,000 rpm for 6 to 7 h at 4°C. The fractions containing the replicating intermediates of monomers were pooled and precipitated with ethanol. The sucrose was removed from the DNA solution by extraction performed twice with phenol. Restriction enzyme digestions were carried out at 25°C, and the reactions were stopped by adding *N*-lauroyl sarcosine to a final concentration of 0.5%.

Electrophoresis. Two-dimensional gel electrophoresis was performed as described by Brewer and Fangman (4) with some modifications. The first dimension was 0.4% agarose in 0.5× TBE (1× TBE contains 89 mM Tris-89 mM boric acid-2 mM EDTA [pH 8.0]) and run at 1 V/cm for 20 h (unless otherwise stated) at 25°C. The second dimension (1% agarose) was carried out at 5 V/cm in  $0.5 \times$  TBE containing 0.3 µg of ethidium bromide per ml at 4°C. In the experiments illustrated in Fig. 6A and 7, the second dimension was performed at 6 V/cm for 12 to 13 h in 1× TBE in a constant-temperature room held at  $-15^{\circ}$ C. In this condition, the electrophoresis buffer was in an equilibrium of solid and liquid phases. The reverse-field gel electrophoresis was performed in 0.8% agarose and 0.5× TBE for 18 to 20 h at 3.5 V/cm. The setting of the pulse controller (Hoefer PC750) was as follows: forward, 150 ms; reverse, 50 ms; and ramp factor, 0.3.

#### RESULTS

The replication intermediates of BPV-1 detected by 2D gel electrophoresis show composite patterns. The structure of DNA replication intermediates can be analyzed by the 2D gel electrophoresis technique developed by Brewer and Fangman (4). In this gel system, the first dimension separates molecules under conditions wherein the friction factor of the different species varies in a simple and unidirectional way with the molecular mass of the species. In the second dimension, which is run in a higher agarose concentration and in the presence of ethidium bromide, the friction factor is much more dependent upon the shape of the species. Figure 1 shows several distinct patterns generated by dif-



FIG. 1. The 2D gel electrophoresis patterns generated by different types of replication intermediates (adapted from Brewer and Fangman [4]). The conditions in the first dimension are designed to separate the molecules largely on the basis of their masses. The second dimension fractionates the molecules on the basis of their shapes. The expected 2D gel electrophoresis patterns are shown below each type of replicating intermediate. (A) The Y arc is generated by molecules with replication forks moving unidirectionally from one end to the other. (B) The bubble arc results from the DNA fragments containing an initiation bubble in the middle. (C) The straight-line pattern is generated from fragments with two replication forks migrating from both ends at about the same speed. (D) The transition pattern from a bubble arc to a Y arc is expected when the origin of replication is asymmetrically located on the fragments.

ferent types of replication intermediates. For example, if after restriction enzyme hydrolysis a fragment is placed with respect to an origin of replication such that the fork moves into the fragment from one direction, a series of the sort called a Y pattern is generated. The theoretical and empirical data which support these interpretations are given in references 4 and 5.

BPV-1 DNA was isolated from unsynchronized growing ID13 cells by the Hirt extraction method described in Materials and Methods. The extract was then digested with the restriction enzymes HindIII and SmaI. This hydrolysis splits the viral genome into two fragments of 6 and 1.9 kb. After 2D gel electrophoresis, the DNA was transferred to nitrocellulose and probed with a sequence which hybridizes only with the 6-kb fragment (Fig. 2). In the figures, the probes utilized are diagrammatically shown as boldface lines. The major hybridization shows a spot at 6 kb which corresponds to the intact fragment (Fig. 2). Molecular weight standards were run in parallel for the first dimension and used to assign the sizes of spots. (An example of such a standard is shown in Fig. 7, where lengths in the first dimension were critical.) The replication patterns detected clearly show two overlapping patterns of about equal intensity. One pattern is a line pattern emanating from the 6-kb spot in the arc of linear DNAs. This line (labeled a in Fig. 2) is generated from double Y-shaped molecules (Fig. 1, pattern C). This pattern suggests that this series of molecules has replication forks which migrate from both ends and meet near the middle of the fragment. One can infer that this series could be generated by an origin of replication placed between the *HindIII* and *SmaI* sites in the small fragment, as shown in Fig. 2. The second pattern detected complicates this simple picture, as it is a Y arc which is labeled b in Fig. 2. There are several ways to explain this Y arc. It could



FIG. 2. 2D gel analysis of the total BPV-1 DNA from ID13 cells. The BPV-1 DNA from a Hirt supernatant (see Materials and Methods) was digested with *Hind*III and *Sma*I and analyzed by 2D gel electrophoresis. The DNA was then transferred to a nitrocellulose filter and probed with a sequence hybridizing with the 6-kb fragment (boldface line). a, Straight-line (or double Y) pattern; b, Y arc pattern. The numbers indicate the sizes of the linear DNAs. The 8-kb spot is the product of partial digestion (all 8-kb spots in this paper are the linear unit length BPV-1 DNA resulting from partial digestion). The diagram shows our interpretation of how patterns a and b are generated (see text).

reflect a mode of replication different from those represented in the line pattern, for example, a unidirectional mode, or it could result from many nonspecific initiation events occurring over a large portion of the viral genome. Such pseudorandom starts have been suggested to occur during the vegetative growth of BPV-1 in warts, as judged by eye forms detected in packaged virions (7). It is also possible that the Y arc is an artifact of manipulation (28). For example, branch migration might occur only at one end of double Y molecules. However, in multiple repeats of this experiment with different batches of DNA or with digestions of the replication intermediates with other enzymes, such composites were always detected and had about the same half-and-half distribution.

Another possibility, and the one proven correct by our subsequent studies, is that the Y arc results from oligomeric forms of replicating BPV-1 DNA. Such oligomers have been detected by a number of workers [1, 43] and can make up a significant fraction of the total BPV-1 pool. The diagram shown in Fig. 2 illustrates how a Y arc can be generated from dimers (an equivalent argument can be made for any oligomer). If on any given molecule only one of the two origins is used in a single round, the forks traverse the 6-kb HindIII-Smal fragment (indicated by boldface lines) from one end to the other. This traversal produces a Y arc. To test this hypothesis directly, we examined our DNA preparations for oligomers and subsequently utilized successive sucrose gradient fractionation to isolate and examine only the monomer intermediates. The disadvantage of this protocol is that we cannot comment in depth about the replication modes of the oligomers; moreover, the monomers may have alternative modes of replication with non-theta-form intermediates. However, no evidence for these types of intermediate exists, while theta forms have been detected (43). After purification of theta-form intermediates of monomers, the patterns did simplify, and we were able to extend our analysis to the monomer forms without ambiguity.

BPV-1 DNA from transformed mouse cells is composed of a



FIG. 3. The BPV-1 DNA preparation contains highly heterogeneous multimers. To resolve the multimers of BPV-1 DNA, the samples were analyzed by reverse-field gel electrophoresis. The conditions used here can separate circular DNA from 6 up to 100 kb (see Materials and Methods). Lane A, Undigested BPV-1 DNA; lanes B to G, time course of BPV-1 DNA digestion with *Eco*RI (2, 5, 10, 20, 40, and 80 min, respectively); lane H, linear DNA size marker (Bethesda Research Laboratories); lane I, undigested pBS99; lane J, undigested BPV-1. The abbreviations S, N, and L stand for supercoiled (I), nicked (II), and linearized (III) forms of DNA, respectively. The numbers 1, 2, 3, and so on represent monomers, dimers, trimers, and so on, respectively.

highly heterogeneous concatemeric mixture. The forms of the BPV-1 DNA in our Hirt extracts can be separated by reverse-field gel electrophoresis (see Materials and Methods). An example of such a gel is displayed in Fig. 3. The undigested BPV-1 DNA extracted from ID13 cells was resolved into a multiband ladder (Fig. 3, lanes A and J). To determine the structures and sizes of the molecules, the BPV-1 DNA was subjected to EcoRI digestion. BPV-1 contains only one EcoRI site. If the DNA species with lower migrating mobilities than those of monomers have a catenated structure, a partial digestion with EcoRI produces linearized DNA with only monomer size. Alternatively, if those DNAs have a structure of head-to-tail concatemers, a partial EcoRI digest generates linearized DNA with different unit sizes. The time course of an EcoRI digestion which produces partial digests is shown in Fig. 3 (lanes B to G). The complete digest is shown in lane G, where an 8-kb linear DNA is detected. In the time course, new bands are superimposed on top of the uncut pattern, and relative to molecular weight standards shown in lane H, they correspond to a linear BPV-1 monomer, dimer, trimer, and so on (labeled 1L to 5L in Fig. 3). These bands appear after 2 min of digestion (Fig. 3, lane B), and new species of up to five times the linear monomer length are clearly visible. As the hydrolysis proceeds, the slower-migrating bands disappear and the faster linear species accumulate (compare lanes B and F in Fig. 3). These data demonstrate that BPV-1 DNA from ID13 cells exists as head-to-tail concatemers. Moreover, the unit step size increases by intervals of one monomer. We ascribe the supercoiled forms in Fig. 3, lane A, to the darker species in each step, while the nicked circles are the weaker species in the ladder. The relative amounts of monomer, dimer, trimer, etc., are substantiated by the amount of linear DNA found in the early time points of the partial digestion. Thus, the amount of monomer form I supercoils is about equal to that of the form I supercoils of the dimer, while the trimer supercoils are less. These ratios are also found in linear DNAs which appear in the partial digests. It is also evident from the supercoiled ladder that BPV-1 oligomers up to 12-mers are present.

To confirm the sizes of the oligomers, a direct comparison of BPV-1 DNA with a known multimeric plasmid marker, pBS99, was made. In RecBC<sup>-</sup> SbcB<sup>-</sup> strains of *E. coli*, rolling circle amplification of pBR-based vectors occurs (10). The monomer size of pBS99 is 6 kb, and the resolved circular supercoiled forms of the rolling-circle amplification products increase in size by steps of one monomer (J. Bliska, Ph.D. thesis, University of California, Berkeley, 1984). Undigested BPV-1 DNA (Fig. 3, lane J) was fractionated side by side with supercoiled plasmid pBS99 (lane I). A semilogarithmic plot of the relative mobility of the BPV-1 supercoils with the standard shows clearly that the oligomers increase in size by unit length. That is, a monomer, a dimer, a trimer, a tetramer, and so on are observed (data not shown).

Densitometric analysis of these gels shows that about 45% of the BPV-1 DNAs are in oligomeric forms. Such a heterogeneous population of DNA must give overlapping patterns in the 2D gel analysis if the oligomers replicate. As BUdR labeling experiments performed by us (L. Yang, J. Reynolds, and M. Botchan, unpublished data) and others (13) show that oligomers do enter into the pool of replicating molecules (see discussion), a simple Brewer and Fangman analysis cannot be performed on such a mixture. Our screen of different established BPV-1 lines in our laboratory and those obtained from others showed that all such lines contained a significant amount of oligomers, though the ratios were different (data not shown).

To examine our replication intermediates from monomers, BPV-1 DNA was fractionated by sucrose gradient centrifugation. Figure 4A shows the distribution of BPV-1 forms fractionated in the first cycle of centrifugation in a 10 to 30% sucrose gradient. The monomer supercoils peak in fractions 18 through 20, and their replicating intermediates can be detected as the smear of hybridization between the form I and form II positions. The trimer and higher forms are clearly fractionated away from the monomers. However, a significant amount of dimer cross-contaminates the monomer pools. Thus, a second round of sucrose gradient centrifugation on the pool of monomers was performed. Fractions 18 through 24 contain mainly form I monomers (Fig. 4B). Notice that form II (nicked circle) of the monomers migrates more slowly in the gradient than the monomer supercoiled form does. Thus, the nicked circle monomer peaks in fraction 26. Similarly, the nicked circle of the dimer runs more slowly than does form I of the dimer. Thus, form II of the dimer cross-contaminates in fractions 20 to 22. However, the replication intermediates of the dimers are present in fractions 12 to 14, which do not cross-contaminate the monomers' supercoils. The migration of these forms is as predicted from previous centrifugation experiments with supercoiled molecules and oligomers. Moreover, we surmise that the heterogeneous forms migrating between the form I



FIG. 4. Detection of BPV-1 DNA forms in a sucrose gradient centrifugation analysis. About 300  $\mu$ l of BPV-1 DNA was loaded onto a 10-ml linear sucrose gradient (10 to 30%) in 1 M NaCl-20 mM Tris (pH 7.7)-1 mM EDTA. The centrifugation was carried out in a Beckman SW41 rotor for 6 to 7 h at 36,000 rpm. Fractions (300  $\mu$ l each) were collected, and 3- $\mu$ l samples from each fraction were analyzed by electrophoresis through a 0.8% agarose gel in 0.5× TBE. After electrophoresis, the samples were transferred to nitrocellulose. The blot was probed with intact BPV-1 DNA. (A) First round of centrifugation. The pooled fractions containing monomers and their replicating intermediates were loaded onto the sucrose gradient for the second round of centrifugation. (B) Profile of the second sucrose gradient centrifugation. The replicating intermediates of monomers (fractions 18 to 24) were isolated from the second gradient (called pooled BPV-1 DNA) and used for 2D gel mapping analysis thereafter. The arrows indicate the direction of the sedimentation.

and form II species corresponds to the replication intermediates of BPV-1 by analogy to what has been found for other supercoiled intermediates, such as SV40 DNA replication intermediates (12). When the pooled monomer material from the second sucrose gradient was analyzed by 2D gel electrophoresis, only single patterns were detected. For example, the Y arc illustrated as curve b in Fig. 2 was preferentially lost. The subsequent data shown in Fig. 5 to 7 were obtained with such material.

Mapping the origin of DNA replication. The results presented in Fig. 2 indicate that the monomer forms of BPV-1 replicate bidirectionally and that the forks meet and terminate near the center of the 6-kb HindIII-Smal fragment. To extend and test this prediction, the monomer pools collected as described above were subjected to other 2D gel analyses. DNA was hydrolyzed with both BamHI and KpnI; this digestion produces two BPV-1 fragments (Fig. 5). The hydrolyzed samples were fractionated by 2D gel electrophoresis, and after transfer to nitrocellulose the blot was hybridized with the specific 6.94-kb probe shown in Fig. 5. The results shown in Fig. 5 demonstrate that a bubble arc arises from the 6.94-kb spot. This arc is created by an initiation site within the fragment proceeding towards its ends (Fig. 1, pattern B). The open arrow in the figure points to the end of a Y arc pattern. We note that the bubble arc in Fig. 5 extends past the downward arm of the Y arc. This is because, as illustrated in the cartoon in Fig. 1, nearly completed bubble forms (which resemble form II molecules) migrate in the first dimension more slowly than do any of the nearly completed Y forms (which resemble linear molecules). This creates a displacement of one curve with respect to the other. Thus, if the bubble initiates close to but not precisely at the center of a DNA fragment, the bubble arc converts to a Y arc. The data shown in Fig. 5 thus suggest that the bubble does not initiate precisely in the center of the 6.94-kb fragment. Rather, they indicate that the initiation site is asymmetrically placed and that at later stages in the replication elongation process the molecules which form the bubble arc are converted into Y-shaped molecules. Some of the molecules in this line may arise from artificial breaks in the bubble forms, a common artifact in these experiments (28). Moreover, the separation of oligomers is not absolute and therefore trace levels of Y arc due to these forms are noticed on longer exposures. These experiments, taken together with the previous data, show that BPV-1 initiates its DNA replication within the small *Hind*III-to-*Sma*I fragment.

To increase the mapping resolution, smaller DNA frag-



FIG. 5. 2D gel analysis of the larger BamHI-KpnI fragment. The pooled BPV-1 DNA from a sucrose gradient was digested with BamHI and KpnI and fractionated by 2D gel electrophoresis. The boldface line in the upper part of the diagram on the right shows the 6.94-kb fragment detected by hybridization. The lower part of the diagram illustrates how the 2D gel pattern on the left was generated. The open arrow marks the late part of the Y arc, where the molecules have the shape of a large Y or a highly asymmetric double Y.



FIG. 6. The replication initiation site is closer to PMS-1 than to PMS-2. (A) The pooled BPV-1 DNA from a sucrose gradient was digested with *Hind*III and *Eco*RI and subjected to 2D gel electrophoresis. The probe used was the 3.1-kb *Hind*III-*Eco*RI fragment (boldface line in the diagram). (B) The pooled BPV-1 DNA was cleaved with *Hpa*1 and *Kpn*I. The resulting 2D blot was hybridized with the 3.45-kb *Hpa*I-*Kpn*I fragment (boldface line in the diagram).

ments containing the replication initiation bubble were examined. The strategy used here was transition point mapping. The restriction sites were selected so as to place the putative initiation bubble asymmetrically in a fragment, so that a transition from the bubble arc to a Y arc would occur (Fig. 1D). The early part of a bubble arc can be readily distinguished from the early part of the Y arc, as it has a steeper slope emanating from the spot which detects the duplex fragment. The extent of replication of different molecules along the Y arc has been determined by direct electron microscopic examination (5). For example, the molecules at the turning point in the arc (the minimal mobility point in the second dimension) have been shown to contain molecules which are 50% replicated. In order to increase our resolution and to minimize artifacts due to branch migration of the nascent strands from the early replicating forms, the gels were run in a  $-15^{\circ}$ C cold room as described in Materials and Methods.

One example of such transition point mapping is shown in Fig. 6. Pooled monomeric BPV-1 DNA and its replicating intermediates were digested with HindIII and EcoRI and subjected to 2D gel electrophoresis. The blot was hybridized with the 3.1-kb HindIII-EcoRI fragment (boldface line in Fig. 6A). The autoradiograph is shown in Fig. 6A. A bubble arc arises from the 3.1-kb spot and then converts sharply to a Y arc (compare pattern D in Fig. 1). The length of the newly replicated DNA strand at the transition point was estimated to be about 1.36 kb (or 44% of the 3.1-kb fragment). The inflection point in the Y arc is taken as the 50%-replicated position and the 3-kb spot represents the 0%-replicated molecules. To a first approximation, the extent of replication between 0 and 50% can be found by measuring the point of transition and assuming a close-tolinear relationship between the migration distance and the extent of replication. For example, molecules located halfway between the inflection at the Y arc maximum and the



FIG. 7. Fine mapping of the BPV-1 origin of replication. The pooled BPV-1 DNA from a sucrose gradient was cleaved with MluI and EcoRI and subjected to 2D gel analysis. The first dimension was run at 1 V/cm for 35 h at 25°C. The blot was probed with the 2.7-kb MluI-EcoRI fragment (indicated by the boldface line). The size markers are the references for the first dimension. The diagram illustrates the replicating intermediates in this transition pattern.

unit length spot have been shown to be 25% replicated (5). Since replication is bidirectional, the initiation bubble is located in the middle of the duplicated region (i.e., about 0.68 kb away from one end of the fragment). This result does not tell us on which side of the fragment the initiation bubble is asymmetrically located. It could be closer to the HindIII site, or it could be nearer the EcoRI region. To distinguish between these two possibilities, the replication intermediates generated in the 3.45-kb restriction fragment from HpaI to KpnI (labeled by the boldface line in Fig. 6B) were analyzed by 2D gel electrophoresis. If initiation started at about 0.68 kb from the EcoRI site (diagrams in Fig. 6), a pattern of a bubble arc would be expected. The result is shown in Fig. 6B. A Y arc can be seen clearly, but no sign of a bubble arc can be observed. These data suggest that the replication fork moves unidirectionally throughout this piece of DNA; therefore, the replication starting site is asymmetrically located at a site closer to PMS-1 about 680 bp from the HindIII site.

It is interesting that this calculation localizes the replication starting site at a distance from the 3' border of the PMS-1 region (29, 31). To confirm this result and to increase resolution further, an even smaller fragment than that used in the experiment whose results are shown in Fig. 6A was analyzed. The pooled BPV-1 DNA was treated with EcoRI and MluI and subjected to 2D gel electrophoresis. A transition pattern was again observed (Fig. 7). The transition from the bubble arc to the Y arc was measured about 760 bp away from the MluI site (i.e., about 28% of this 2.7-kb fragment). Since replication is bidirectional, the initiation site is therefore localized about 380 bp away from the MluI site. The markers used for measuring the spot for the transition are shown below the 2D pattern. The autoradiograms do not allow for a precise measurement of the transition points. Thus, an uncertainty of between 100 and 200 bp in the length measurement is found; therefore, the origin can be mapped to approximately  $\pm 100$  bp. The data in Fig. 6A lead to an estimate that the origin of DNA replication is located near BPV position 7640, and the data in Fig. 7 measure the origin at 7731. Each of these estimates is thus consistent with the other and places the origin external to the BPV-1 ClaI C fragment, which was defined as PMS-1.



FIG. 8. The diagram shows the relative location of PMS-1 defined by genetic studies and the viral replication initiation region mapped by 2D gel electrophoresis. The *Cla*I C fragment (defined by BPV-1 coordinates 6834 to 7477) was shown by Lusky and Botchan (29) and Waldeck et al. (43) to provide sufficient information for plasmid maintenance. Moreover, linker insertion analysis defined, in a transient amplification assay, those sequences in the open box labeled PMS-1 as critical for replication. The E2RE1 enhancer was defined by Spalholz et al. (38). The four E2-binding sites, 7 to 10, are shown in the E2RE1 region.

# DISCUSSION

The important new finding reported here is that the actual initiation of BPV-1 DNA replication, as determined by 2D gel electrophoresis, lies 3' to the region defined as sufficient for plasmid maintenance (Fig. 8). Lusky and Botchan (29) and Waldeck et al. (43) were able to show persistence of plasmids that contained a genetic marker linked to the BPV-1 ClaI C fragment in BPV-1-transformed cells. Moreover, in a transient replication assay Lusky and Botchan (31) did not find any critical need for sequences in the BPV-1 URR that lies between the MluI site at BPV position 7352 and the HpaI site at map position 7947/1 (Fig. 8). Several interesting possibilities need to be explored. Although sufficient information for replication exists within the ClaI C fragment, this information may be redundant, as the most potent initiation sequences in the context of the entire viral genome lie close to BPV-1 position 7730. Furthermore, information within the origin region (that is, the region around position 7730) may not be sufficient, and initiation may require information present in the upstream region for effective function. Thus, the PMS-1 region may serve as a facilitator (or enhancer) for replication at a distance in much the same way as the  $\gamma$  origin does for R6K plasmids in E. coli (35). Similarly, the amplification control element 3 described for the Drosophila chorion gene amplification system is critical for origin function and can serve as an origin itself, but initiation as measured by 2D gel mapping occurs most frequently in downstream sites (17). The EBV origin of replication (ori-p) is also a complex structure stretching over a distance of 1 kb (33). It is unclear why both clusters of EBNA-1-binding sites which define the boundaries of ori-p are necessary for origin function. One view of this complexity is that in the multipartite structures shared by all of these origins of replication, the flanking sequences simply serve an enhancers for biochemical events that take place at the origin core. This role would be analogous to the role frequently ascribed to enhancers of transcription. For example, protein-DNA interaction within the upstream regions may stabilize protein-DNA (or protein-protein) interaction at the downstream initiation site. In the BPV-1 and chorion gene cases, however, it cannot explain why the isolated putative enhancers have origin function themselves. This argument tends to support the notion that the information is redundant. Neither point of view is, of course, mutually exclusive of the other. Thus, redundancy and enhancer interactions may both occur. Finally, it is important to mention that DNA fragments taken out of the context of the intact viral genome may assume properties not necessarily relevant to function in the virus.

Two previous reports provided data that are consistent with a single origin of DNA replication in BPV-1 plasmids which functions bidirectionally. The electron-microscopic data of Waldeck et al. (43) placed the origin of replication within the PMS-1 region. It seems possible to us, however, that their alignments for some of the replication eyes scored were in error. In aligning a series of molecules in which the eyes are not dramatically asymmetric with respect to an end, a choice regarding which end is 5' and which is 3' has to be made. In a small series, such an erroneous choice for just a few molecules can distort the mapping and the overall margin of error. Waldeck et al. used two different restriction enzyme-generated linear DNAs in order to map the replication bubble, but in both cases the sample set was small (10 molecules). The data was certainly consistent with a placement of the initiation site within the ClaI C fragment. However, with their margin of error in their length measurements (5%) and the intrinsic alignment problems described above, it is conceivable that our data are in fact compatible with theirs. Our data are also consistent with the recent 2D gel mapping studies of Schvartzman et al. (37), who did not attempt to map the origin beyond locating it in a region within the URR.

The method used here to map the origin of BPV-1 replication is limited by two problems. First, the transition from one pattern to another occurs over a region, and one must estimate the break point. Second, it is assumed that the mobility of the replication intermediates in the first dimension is the same as that of the linear duplexed markers. For the almost completed replicating molecules, this is clearly the case, as the final forms intersect at two times the size of the analyzed restriction fragment (Fig. 7). For other points in the curve of the Y arc, however, such assignments are only approximate. For example, molecules located halfway up the Y arc have been shown to be about one-fourth replicated (5). We have presented a quantitative evaluation for two independent 2D gel electrophoresis patterns, and in both cases the data are consistent. The qualitative interpretations of the Brewer and Fangman analysis are unequivocal. The qualitative data shown in Fig. 7 prove that the replication initiation site in the BPV-1 DNA must be 3' to the MluI site. This is because if DNA replication initiated in the PMS-1 region, only a Y arc rather than a transition curve would be observed for a fragment 3' to this element. These points are substantiated by the patterns shown in Fig. 6. Furthermore, if the origin of replication were placed close to the HindIII site at BPV-1 position 6959 (Fig. 8), the line pattern seen in Fig. 2 would be impossible. This last point was first brought to our attention by B. Brewer in her analysis of our preliminary data.

The 2D gel mapping experiments reported here place the BPV-1 origin of DNA replication within the region of the URR initially described as the E2-responsive enhancer element 1 (E2RE1) of the viral DNA (38). Spalholz et al. (38) mapped two discrete regions within the URR that could activate the SV40 promoter provided that the E2 protein was present as a *trans*-acting factor. The other region, called E2RE2, overlaps with PMS-1. The E2RE1 region contains four high-affinity binding sites for E2 and has the greatest

density of such sites in the viral genome (27). Consequently, it is the most efficient E2-responsive enhancer element (2, 16, 34, 39). E2RE2 contains three E2-binding sites, one of which is a high-affinity site. Moreover, the ClaI C fragment contains four E2-binding sites. Therefore, both PMS-1 and the region containing the DNA replication initiation site have E2-binding sites. We have recently found that the BPV-1 E1 protein forms a specific complex with the BPV-1-encoded E2 protein. Moreover, the complex binds to E2RE1 with high affinity and binds with lower affinity to the upstream sites within the ClaI C fragment (I. Mohr, R. Clark, S. Sun, E. Androphy, P. MacPherson, and M. Botchan, submitted for publication). The role this protein complex plays in the initiation of BPV-1 DNA replication is under investigation. Our present model is that an E1-E2 protein complex is required to assemble the replication initiation machinery.

In previous studies, we had discounted a direct role for the E2 gene product in BPV-1 DNA replication because the Bal-15 mutant (which deletes all BPV-1 open reading frames 3' to the E1 open reading frame, including E2) appeared to persist in stable or transient assays (29). We have not been able to reproduce those results, and we now assume that those results were due to plasmid persistence or repair synthesis and not semiconservative replication (M. Lusky, L. Thorner, and M. Botchan, unpublished observations). However, we cannot eliminate the possibility that our current C127 cells are somehow modified in cellular components which once were present and could substitute for E2. In any case, all present genetic evidence indicate that E2 is absolutely required for BPV-1 plasmid maintenance. Moreover, Stenlund and Ustav have recently shown that a fragment containing the URR does not replicate transiently when cotransfected with an E1 expression vector but does so when cotransfected with both an E1 and an E2 expression vector. Thus, the in vivo requirement for E2 expression consistent with previous genetic studies (32) cannot be ascribed to merely the activation of the expression of the E1 protein (A. Stenlund and M. Ustav, personal communication).

We have found that the plasmid DNA forms of BPV-1 present in the ID13 cell line as well as those in other established lines, including VI216 and WHZ-10, are highly heterogeneous and consist of head-to-tail concatemers. Similar observations have been reported by others (1, 37, 43). These oligomers replicate, and their replication intermediates become superimposed upon those of the monomers in 2D gel analysis. Thus, for an oligomer, only one of the many potential origins initiates in a given round of replication. Our interpretation is similar to that recently reported by Schvartzman et al. (37). Conceptually, if all origins of replication on an oligomer were utilized at once, then the 2D gel electrophoresis patterns of those species would be the same as those of the monomers. In contrast, if only one of the many potential origins of replication functions in any given round of replication for a given oligomeric molecule, multiple 2D gel patterns are superimposed (Fig. 2). If the rate-limiting step in BPV-1 replication is initiation, this necessarily follows. SV40, for example, requires 20 to 30 min to complete the replication of one monomer in vivo, vet one would predict only 1 to 2 min for completion of replication given the elongation rate of the replication complex (26). Therefore, once replication initiates, elongation proceeds rapidly and results in termination before a second initiation event can occur. Thus, for SV40 evolutionary variants such as ev-1103, which has nine SV40 origins of replication, only molecules with a single bubble are detected (6). In addition,

as DNA replication forks move forward, presumably driven by a helicase, a highly positively supercoiled DNA domain is generated ahead of the forks (44). These transient positive supercoils make the assembly of an initiation complex at the other origins difficult.

These arguments, however, also leave open the possibility that once replication initiates on a given molecule, negative regulation by another mechanism also interferes with subsequent initiation. Along these lines, it is interesting to point out that bromodeoxyuridine labeling occurs for both the monomers and oligomers. In density shift experiments the fraction of oligomers in all of the peaks is equivalent to the fraction of oligomers in the starting material (J. Reynolds, L. Yang, and M. Botchan, unpublished observations). If entry of molecules into the replication pool were random and chosen simply by a stochastic interaction between transacting factors and origins, this would not occur. The multimers would be preferentially labeled, as those molecules have more origin sites. The mechanisms that create the concatemerized molecules are therefore of interest. These head-to-tail circles may arise by errors in termination during a round of replication as they do for SV40 (14). However, this model would predict that they initially form in an integral series divisible by two (a monomer to a dimer, dimer to a tetramer, and so on). Subsequent recombination could perhaps form the odd-numbered members. We suspect, however, that the oligomers occur sporadically and do not accumulate in any systematic way. This is because the patterns are characteristic for a given cell line. Recombination between monomers could also produce a ladder with an apparent exponential decay such as that shown in Fig. 3 for the BPV-1 oligomers. However, we point out that the characteristic patterns in the 2D gel analysis blots (5) expected for recombination intermediates were not detected. This supports the argument that recombination events are rare. Finally, it will be important to explore the possibility that the high-molecular-weight oligomers (larger than pentamers) arise through a rolling-circle mode of replication which has been suggested by Burnett et al. (7). This mode of replication could be significant for any amplification of BPV-1 plasmids which might occur in a subpopulation of cells.

### ACKNOWLEDGMENTS

We are grateful to B. Brewer for helpful discussions and I. Mohr for critical reading of the manuscript.

This work was supported by Public Health Service grant CA42414 to M.B. from the National Institutes of Health. L.Y. is a recipient of a Damon Runyon-Walter Winchell Cancer Research Fund postdoctoral fellowship.

## LITERATURE CITED

- Allshire, R., and C. Bostock. 1986. Structure of bovine papilloma virus type I DNA in a transformed mouse cell line. J. Mol. Biol. 188:1–13.
- 2. Androphy, E., D. Lowy, and J. Schiller. 1987. Bovine papilloma virus E2 trans-acting gene product binds to specific sites in papilloma virus DNA. Nature (London) 325:70–73.
- 3. Berg, L., M. Lusky, A. Stenlund, and M. R. Botchan. 1986. Repression of bovine papilloma virus replication is mediated by a virally encoded trans-acting factor. Cell 46:753–762.
- 4. Brewer, B., and W. Fangman. 1987. The localization of replication origins on ARS plasmids in S. cerevisiae. Cell 51:463-471.
- Brewer, B., E. Sena, and W. Fangman. 1988. Analysis of replication intermediates by two dimensional agarose gel electrophoresis. Cancer Cells 6:229-234.
- 6. Brockman, W., T. Lee, and D. Nathans. 1974. Characterization

of cloned evolutionary variants of simian virus 40. Cold Spring Harbor Symp. Quant. Biol. **39:**119–127.

- 7. Burnett, S., J. Zabielski, J. Moreno-Lopez, and U. Pettersson. 1989. Evidence for multiple vegetative DNA replication origins and alternative replication mechanisms of bovine papilloma virus type I. J. Mol. Biol. 206:239-244.
- Challberg, M., and T. Kelly. 1989. Animal virus DNA replication. Annu. Rev. Biochem. 58:671–717.
- Clertant, P., and I. Seif. 1984. A common function for polyoma virus large-T and papillomavirus E1 proteins? Nature (London) 311:276–279.
- Cohen, A., and A. J. Clark. 1986. Synthesis of linear plasmid multimers in *Escherichia coli* K-12. J. Bacteriol. 167:327-335.
- 11. Crosa, J. 1980. Three origins of replication are active in vivo in the R plasmid RSF1040. J. Biol. Chem. 255:11075–11077.
- 12. **DePamphilis, M., and M. Bradley.** 1986. Replication of SV40 and polyoma virus chromosomes, p. 99–246. *In N. Salzman* (ed.), The Papovaviridae. Plenum Publishing Corp., New York.
- 13. Gilbert, D., and S. Cohen. 1987. Bovine papilloma virus plasmids replicate randomly in mouse fibroblasts through S phase of the cell cycle. Cell 50:59–68.
- 14. Goff, S. P., and P. Berg. 1977. Structure and formation of circular dimers of simian virus 40 DNA. J. Virol. 24:295–302.
- Hand, R. 1978. Eukaryotic DNA: organization of the genome for replication. Cell 15:317–325.
- Hawley-Nelson, P., E. Androphy, D. Lowy, and J. Schiller. 1988. The specific DNA recognition sequence of the bovine papilloma virus E2 protein is an E2-dependent enhancer. EMBO J. 7:525– 531.
- 17. Heck, M., and A. Spradling. 1990. Multiple replication origins are used during Drosophila chorion gene amplification. J. Cell Biol. 110:903-914.
- Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26:365-369.
- Huberman, J., L. Spotila, K. Nawotka, S. El-Assouli, and L. Davis. 1987. The in vivo replication origin of the yeast 2μm plasmid. Cell 51:473-481.
- Kearsey, S. 1986. Replication origins in yeast chromosomes. Bioessays 4:157-161.
- Kolter, R., and D. Helinski. 1982. Plasmid R6K DNA replication. II. Direct nucleotide sequence repeats are required for an active gamma origin. J. Mol. Biol. 161:45-56.
- 22. Lambert, P., C. Baker, and P. Howley. 1988. The genetics of bovine papilloma virus type 1. Annu. Rev. Genet. 22:235-258.
- 23. Lancaster, W. D., and C. Olson. 1982. Animal papillomaviruses. Microbiol. Rev. 46:191-207.
- 24. Laskey, R., and R. Harland. 1981. Replication origins in the eukaryotic chromosome. Cell 24:283–284.
- 25. Law, M.-F., D. R. Lowy, I. Dvoretzky, and P. M. Howley. 1981. Mouse cells transformed by bovine papillomavirus contain only extrachromosomal viral DNA sequences. Proc. Natl. Acad. Sci. USA 78:2727-2731.
- Levine, A., H. Kang, and F. Billheimer. 1970. DNA replication in SV40 infected cells. J. Mol. Biol. 50:549–568.
- 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.

- Linskens, M., and J. Huberman. 1990. Ambiguities in results obtained with 2D gel replicon mapping techniques. Nucleic Acids Res. 18:647-652.
- Lusky, M., and M. R. Botchan. 1984. Characterization of the bovine papilloma virus plasmid maintenance sequences. Cell 36:391-401.
- Lusky, M., and M. R. Botchan. 1986. A bovine papillomavirus type 1-encoded modulator function is dispensable for transient viral replication but is required for establishment of the stable plasmid state. J. Virol. 60:729-742.
- Lusky, M., and M. R. Botchan. 1986. Transient replication of bovine papilloma virus type 1 plasmids: cis and trans requirements. Proc. Natl. Acad. Sci. USA 83:3609-3613.
- Maio, D., and J. Settleman. 1988. Bovine papilloma virus mutant temperature defective for transformation, replication and transactivation. EMBO J. 7:1197–1204.
- Mecsas, J., and B. Sugden. 1987. Replication of plasmids derived from bovine papilloma virus type I and Epstein-Barr virus in cells in culture. Annu. Rev. Cell Biol. 3:87–108.
- Moskaluk, C., and D. Bastia. 1987. The E2 'gene' of bovine papilloma virus encoded an enhancer binding protein. Proc. Natl. Acad. Sci. USA 84:1215-1218.
- Mukherjee, S. 1988. Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. Cell 52:375– 383.
- Prives, C. 1990. The replication functions of SV40 T antigen are regulated by phosphorylation. Cell 61:735–738.
- 37. Schvartzman, J. B., S. Adolph, L. Martín-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.
- Spalholz, B. A., P. F. Lambert, C. L. Yee, and P. M. Howley. 1987. Bovine papillomavirus transcriptional regulation: localization of the E2-responsive elements of the long control region. J. Virol. 61:2128–2137.
- Spalholz, B. A., Y. C. Yang, and P. M. Howley. 1985. Transactivation of a bovine papilloma virus transcriptional regulatory element by the E2 gene product. Cell 42:183–191.
- 40. Stillman, B. 1989. Initiation of eukaryotic DNA replication in vitro. Annu. Rev. Cell Biol. 5:197–245.
- 40a.Sun, S., L. Thorner, M. Lentz, P. MacPherson, and M. Botchan. 1990. Identification of a 68-kilodalton nuclear ATP-binding phosphoprotein encoded by bovine papillomavirus type 1. J. Virol. 64:5093-5105.
- 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.
- Umek, R., M. Lindkens, D. Kowalski, and J. Huberman. 1989. New beginnings in studies of eukaryotic DNA replication origins. Biochim. Biophys. Acta 1007:1–14.
- Waldeck, W., F. Rosl, and H. Zentgraf. 1984. Origin of replication in episomal bovine papilloma virus type I DNA isolated from transformed cells. EMBO J. 3:2173–2178.
- 44. Yang, L., C. Jessee, K. Lau, H. Zhang, and L. Liu. 1989. Template supercoiling during ATP-dependent DNA helix tracking: studies with SV40 large tumor antigen. Proc. Natl. Acad. Sci. USA 86:6121-6125.