The EMBO Journal vol.15 no.18 pp.5085–5092, 1996

The initiator protein E1 binds to the bovine papillomavirus origin of replication as a trimeric ring-like structure

Juhan Sedman and Arne Stenlund[1]

Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA

[1]Corresponding author

The replication initiator protein E1 binds to the origin of replication of bovine papillomavirus in several forms. E1 can bind to its recognition sequence as a monomer together with the viral transcription factor E2, or as a trimeric E1 complex. The trimerization of E1 is mediated by the sequence-specific binding of E1 to DNA, and results in an E1 complex that is linked topologically to the DNA because the three molecules of E1 form a ring-like structure that encircles the DNA. These results demonstrate that E1 utilizes unusual mechanisms for sequence-specific binding to DNA and for the generation of a structure that encircles the DNA. We believe that these forms of E1 bound to the origin of replication represent intermediates in a transition in the function of E1, from a sequence-specific origin of replication recognition protein to a form of E1 that is competent for the initiation of viral DNA replication.

Keywords: DNA binding/DNA replication/initiator/papillomavirus

Introduction

Viral initiator proteins serve several different functions. Through sequence-specific DNA binding they recognize the origin of replication (ori) within the viral genome and mark the region of replication initiation. They allow the initiation of DNA synthesis through localized melting of the DNA helix, and they may also serve as replication helicases that travel in front of the DNA polymerases, unwinding the duplex DNA (for reviews on SV40 large T-antigen in this capacity, see Borowiec *et al.*, 1990; Fanning and Knippers, 1992). This transition, from a site-specific DNA binding protein to a helicase that travels in front of the replication fork, is poorly understood, but it appears that changes in the biochemical activities of the initiator are required. A likely possibility is that the activities of the initiator protein are altered by multimerization.

The bovine papillomavirus (BPV) genome encodes two proteins that are essential for viral DNA replication (Ustav and Stenlund, 1991). The E1 protein is a 72 kDa nuclear protein with a sequence-specific DNA binding activity that serves to recognize the ori, has DNA-dependent ATPase activity and can function as a DNA helicase (Ustav *et al.*, 1991; Wilson and Ludes-Meyers, 1991; Yang *et al.*, 1991; Holt *et al.*, 1993; Seo *et al.*, 1993a; Thorner *et al.*, 1993; Yang *et al.*, 1993). The E2 protein functions as a sequence-specific activator of transcription (Spalholz *et al.*, 1985; Androphy *et al.*, 1987; McBride *et al.*, 1991), but is also required for viral DNA replication. The minimal ori consists of an ~60 bp-long sequence and contains three elements: an A+T-rich region, a binding site for the E1 protein and a binding site for the E2 protein, all of which are required for replication (Ustav *et al.*, 1991, 1993; Yang *et al.*, 1991; Holt *et al.*, 1993; Seo *et al.*, 1993a; Gillette *et al.*, 1994; Holt and Wilson, 1995).

The initiator E1 is capable of binding to the ori in at least two different forms. By virtue of a protein–protein interaction between the E1 and E2 proteins, E1 can bind cooperatively with E2 to form an E1–E2–ori complex when binding sites for both proteins are present (Mohr *et al.*, 1990; Blitz and Laiminis, 1991; Lusky and Fontane, 1991; Yang *et al.*, 1991; Seo *et al.*, 1993b; Gillette *et al.*, 1994; Le Moal *et al.*, 1994; Sedman and Stenlund, 1995). Binding in this form is highly sequence specific, and genetic evidence demonstrates that the ability to form this complex is essential for replication *in vivo* (Sedman and Stenlund, 1995). In fact, an E2 protein from a different papillomavirus that fails to interact with BPV E1 is unable to support replication together with BPV E1. At higher concentrations E1 alone can also bind to the ori to form an E1–ori complex (Ustav *et al.*, 1991; Wilson and Ludes-Meyers, 1991; Yang *et al.*, 1991; Seo *et al.*, 1993b; Lusky *et al.*, 1994; Sedman and Stenlund, 1995). This complex forms with relatively high affinity but shows low sequence specificity (Sedman and Stenlund, 1995). Thus an essential function for the transcription factor E2 appears to be to interact physically with E1 and to increase the specificity of binding of E1 to the ori. Based on these observations, we suggested previously that a major function of E2 is to serve as a specificity factor for the binding of E1 to the BPV ori (Sedman and Stenlund, 1995).

In this manuscript we have analyzed the E1-containing complexes to determine the stoichiometry of binding and how E1 recognizes the ori in the two different complexes. We demonstrate here that E1 can bind to the ori as a monomer, together with a dimer of E2 (E1–E2–ori complex), and that E1 at higher concentrations can bind to the ori as a trimer (E1–ori complex) that encircles the DNA helix. Our results demonstrate that the binding of E1 in the two complexes is highly related, and we propose that the binding of E1 in these different forms represents intermediates in the assembly of a replication-competent initiator complex on the ori.

Results

E1 binds to the BPV ori as a trimer by itself and as a monomer together with E2

To determine the stoichiometry of binding for the E1–ori and E1–E2–ori complexes, we ascertained the molecular

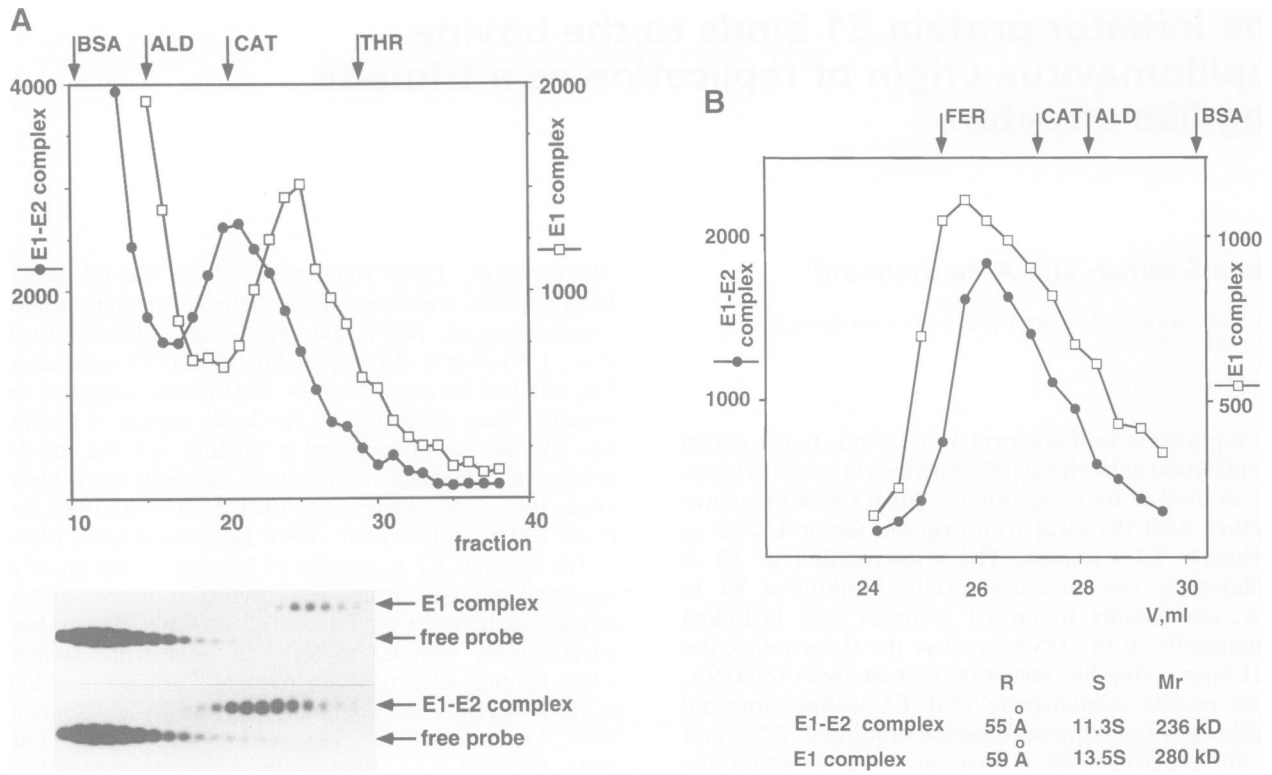
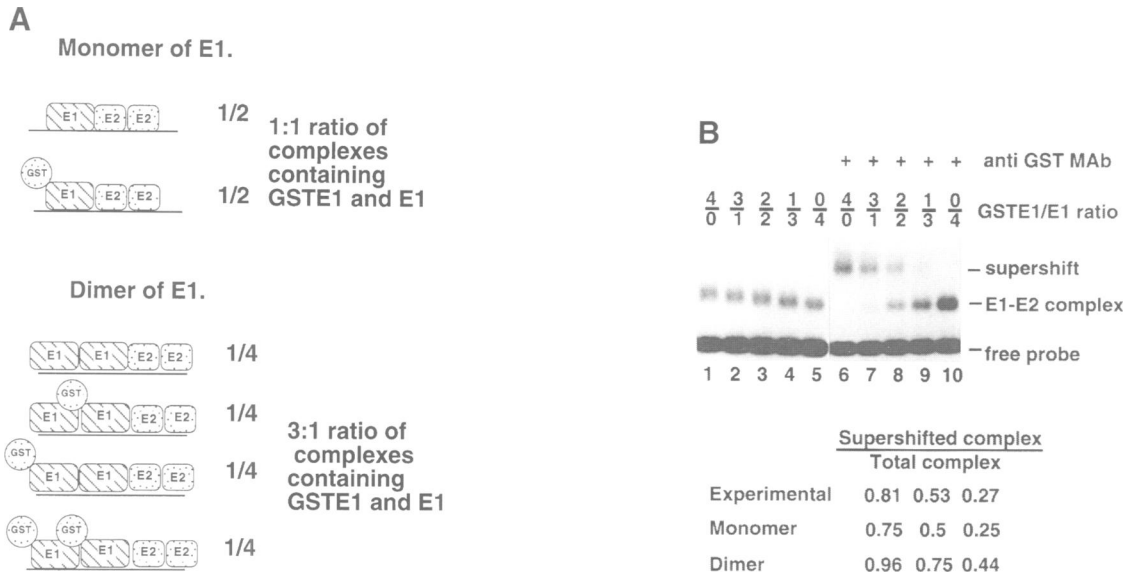


Fig. 1. Molecular mass determination of E1–ori and E1–E2–ori complexes using a gel filtration and sedimentation analysis. E1–ori and E1–E2–ori complexes were generated and crosslinked with glutaraldehyde and analyzed by glycerol gradient sedimentation (**A**) and a gel filtration analysis on Sephacryl S400 (**B**). The standard markers used and their migration are indicated by the arrows at the top of the graphs. The lower part of (A) shows the analysis of the fractions from the glycerol gradient by agarose gel electrophoresis to detect the respective complexes. A summary of the data obtained and the calculated molecular weights is shown in the lower part of (B). The estimated errors represent the standard deviation from three separate determinations of sedimentation coefficients and Stoke's radii.

masses of these complexes. We utilized a combination of glycerol gradient centrifugation and gel filtration to measure the sedimentation values and Stoke's radius for both the E1–E2–ori and the E1–ori complexes (Li and Desiderio, 1993). The complexes were generated by binding E1 and E2 to an origin probe, followed by crosslinking by glutaraldehyde, as described in Materials and methods. Fractions from the gradients were collected and the radioactive peaks identified. Samples were also loaded onto gels, and fractions containing the two complexes identified. The results from the sedimentation analysis and comparison with the sedimentation of marker proteins (see Figure 1A) demonstrated that the E1–ori complex sedimented at 13.5 ± 0.5S and that the E1–E2–ori complex sedimented slightly more slowly at 11.3 ± 0.4S. The Stoke's radii were calculated from gel filtration (Figure 1B) to be 59 ± 4 and 55 ± 2 Å for the E1–ori and the E1–E2–ori complexes, respectively. The combined results from these experiments resulted in a calculated molecular mass of 280 ± 34 kDa for the E1–ori complex and 236 ± 19 kDa for the E1–E2–ori complex. These measurements corresponded most closely to a stoichiometry of a monomer of E1 (70 kDa) and a dimer of E2 (96 kDa) in the E1–E2–ori complex, and to a trimer of E1 (210 kDa) in the E1–ori complex, as the molecular mass of the DNA fragment used was 60 kDa. This stoichiometry was unexpected because the binding site for E1 has a 2-fold rotational symmetry and would be expected to bind an even number of E1 molecules.

To verify the stoichiometry of binding in an independent manner, we used a combined gel shift and antibody supershift assay. The basis for this assay is shown schematically in Figure 2A. We used two different variants of the E1 protein [E1 and an N-terminal glutathione *S*-transferase (GST)–E1 fusion]. If E1 binds as a monomer in the E1–E2–ori complex, at equal concentrations of the two variant proteins, one half of the DNA molecules would be bound by variant 1 (E1) and the other half would be bound by variant 2 (GST–E1). If E1 binds as a dimer in the E1–E2–ori complex and the two variants of E1 can associate freely, three-quarters of the complexes would be expected to contain variant 1 (GST–E1). Because the size difference between these variants is relatively small, the different complexes cannot be resolved based on mobility alone (compare Figure 2B, lanes 1–5). Therefore we used an antibody directed against GST to supershift just those complexes containing GST–E1.

E1 and GST–E1 were mixed at the indicated molar ratios, combined with the probe and each sample divided in two. One set of samples was loaded directly onto the gel (Figure 2B, lanes 1–5); to the other set an excess of antibody was added. The fraction of the complex that could be supershifted by the antibody was compared with the sum of total complex formed. When GST–E1 was used exclusively, all the complex could be supershifted by the antibody (Figure 2B, lane 6). When E1 was used exclusively, none of the complex could be supershifted by the antibody (Figure 2B, lane 10). At the three different ratios (3:1, 2:2 and 1:3; Figure 2B, lanes 7–9) of GST–E1:E1 protein, the ratio between the supershifted complex



Supershifted complex / Total complex			
Experimental	0.81	0.53	0.27
Monomer	0.75	0.5	0.25
Dimer	0.96	0.75	0.44

Fig. 2. Gelshift analysis to determine the stoichiometry of binding in the E1–E2–ori complex. (**A**) A schematic figure showing the expected distribution of GST–E1- or E1-containing complexes if E1 binds as a monomer or a dimer and the two forms of E1 are present in equal quantities. (**B**) Supershift analysis of E1–E2–ori complexes. E1 and GST–E1 were used to generate E1–E2–ori complexes. E1 and GST–E1 were added at different ratios, keeping the total activity constant. One-half of each sample was loaded directly onto the gelshift gel. The other half of each sample was incubated with a monoclonal antibody directed against GST. The ratios of supershifted complex over total complex formed were calculated at the different ratios between E1 and GST–E1. The experimental and calculated values for E1 binding as a monomer and a dimer are shown at the bottom of the figure.

and total complex was very close to that predicted for the binding of E1 as a monomer, and significantly different from the predicted value for dimeric binding. A potential caveat associated with this experiment would be if E1 or GST–E1 could exist in a stable dimeric form in solution, or if homo-association would preferentially take place. Therefore we analyzed the sedimentation of E1 and GST–E1 protein in glycerol gradients; both proteins sedimented exclusively as monomers (data not shown). Furthermore, when the E1–ori complex was analyzed in a similar supershift assay, we observed clear evidence of complexes containing both E1 and GST–E1 (data not shown). These results provide an independent measure for the stoichiometry of E1 binding in the E1–E2–ori complex, and support the conclusions of the molecular mass determinations.

High-resolution footprints demonstrate that E1 binds in a similar fashion in the two complexes

To examine in detail how E1 bound to the recognition sequence in the two complexes, we performed high-resolution hydroxyl-radical footprinting on the two different complexes isolated from gelshift gels. The protections observed with the E1–E2–ori complex show a number of interesting features (Figure 3A, lanes 4 and 9, and B). A 4 bp protection can be observed on each strand over the E2 binding site (stippled boxes). These protections correspond to the binding of E2 because they are observed in the absence of E1 (data not shown). Over the imperfect inverted repeat, three different regions, each corresponding to three to four nucleotides, were protected on each strand (filled boxes). These protections were spaced 10–11 nucleotides apart and arranged symmetrically on the two strands. When the protections were projected onto a double helix, these six protections, which we assume resulted from the binding of E1, were all positioned on the same face of the helix (Figure 3B). The protection over the E2 binding site by E2 was also located on the same face. The simplest interpretation of these results is that the monomer of E1 binds on one face of the recognition sequence, interacting with a dimer of E2 that is also bound to the same face (Figure 3D).

The protections observed in the E1–ori complex (Figure 3A, lanes 3 and 8, and C) were easiest to interpret when compared with the protections observed in the E1–E2–ori complex. Firstly, a general protection was observed over the entire palindrome. This feature indicates that on both strands DNA–protein contacts occurred at virtually every position in the recognition sequence. Secondly, this general protection could be roughly divided into two groups: strong protections and weak protections. The strong protections in all cases represented very specific extensions of the protections observed in the E1–E2–ori complex. Each individual protection observed in the E1–E2–ori complex was extended unidirectionally by three to four nucleotides. On the top strand all the protections were extended in the 3′ direction, and on the bottom strand the extension was in the 5′ direction. On a helix model, these strong protections corresponded to a duplication of the protections observed with the monomer of E1, shifted by three nucleotides, adding up to protection of approximately two-thirds of the circumference of the helix (filled boxes). The simplest interpretation of this result is that two molecules of E1 were bound in a virtually identical fashion but shifted three nucleotides relative to each other: one E1 molecule is bound on top of the helix, with another E1 molecule on the front face of the helix.

In addition to the strong protections, weaker but significant protections can be observed between the strong protections (hatched boxes). On the helix model these represent a further extension of the strong protections by an additional three or four nucleotides, completing the

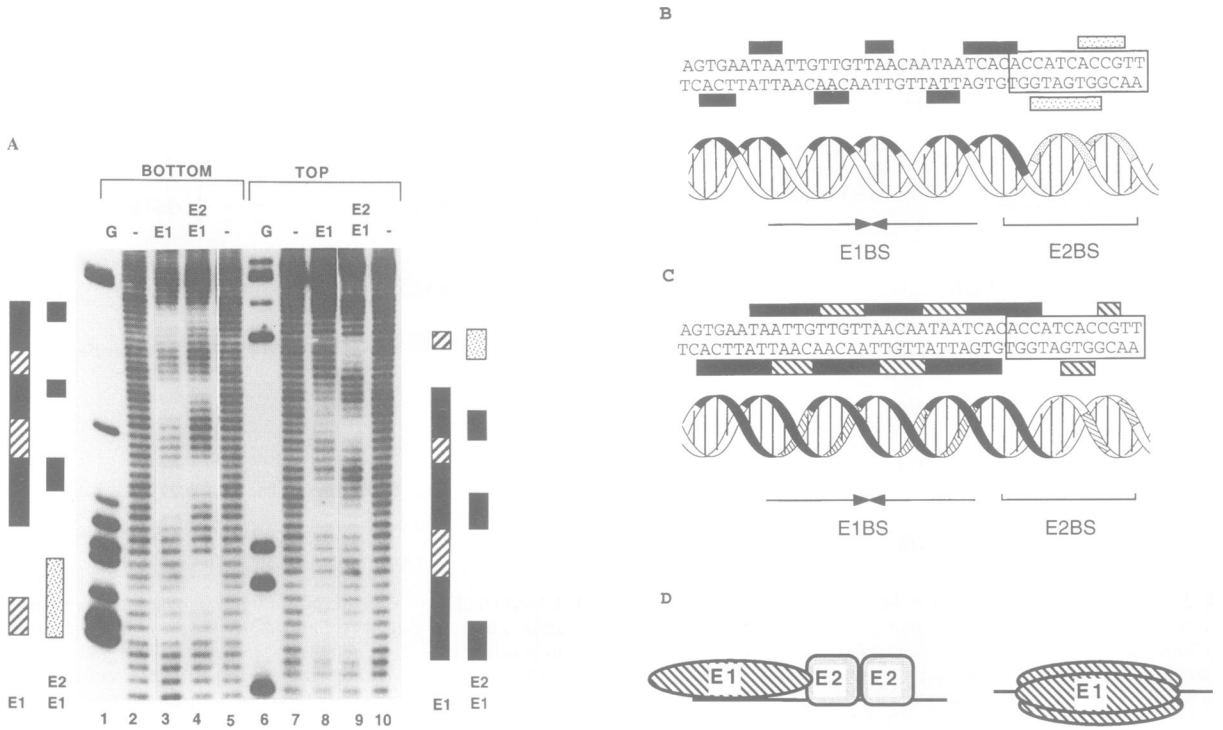


Fig. 3. (**A**) Hydroxyl-radical footprints of the E1–ori and E1–E2–ori complexes isolated after gelshifts. End-labeled origin fragments were used for gel shift assays to generate the E1–ori and E1–E2–ori complexes. The isolated complexes were treated with Fe/EDTA and analyzed on sequencing gels. The lanes labeled G correspond to G-ladders generated by dimethylsulfoxide modification of the respective strands. The protections generated by the E1–E2–ori complex are indicated by solid bars, and the protection over the E2 binding site by stippled bars. The strong protections generated by the E1–ori complex are indicated by solid bars; weak protections are indicated by hatched bars. (**B**) The protected sequences are shown for the E1–E2–ori complex, and the protections are also projected onto a double helix. The protections over the palindrome are represented by solid bars, and the protections by E2 are represented by hatched bars. (**C**) The protected sequences are shown for the E1–ori complex. Strong protections are represented by solid bars, and weak protections by hatched bars. The same symbols were projected onto a double helix model. (**D**) A model for the binding of E1 in the E1–E2–ori complex and in the E1–ori complex.

full circumference of the helix. Two weak protections are also observed over the E2 binding site. Taken together with the data demonstrating that E1 binds as a trimer, these results indicate that two E1 molecules bind tightly on two faces of the helix, and that the third molecule binds less tightly, generating weaker protections on the third face. Binding in this manner would complete a structure of three E1 molecules that encircle the DNA helix. This interpretation would predict that binding sites for at least two molecules of E1 are present within the palindromic sequence, offset by three nucleotides relative to each other. The presence of overlapping binding sites is supported by mutational data which demonstrate that mutations at certain positions in the palindrome affect the formation of only the E1 trimer and not the E1–E2–ori complex (J.Sedman and A.Stenlund, unpublished data). Hydroxyl-radical footprinting using E1 and E2 in solution gave rise to virtually identical protections to those observed with the isolated crosslinked complexes, indicating that the process of crosslinking and isolation of the complex did not in a substantial way change the binding of E1 (data not shown).

Trimerization of E1 is induced by the ori

To determine if a form of E1 crosslinked into a trimer could be detected, we formed the E1–ori complex under standard conditions followed by crosslinking with glutaraldehyde. As controls, one sample was assembled in the absence of ori DNA and one sample was not treated with

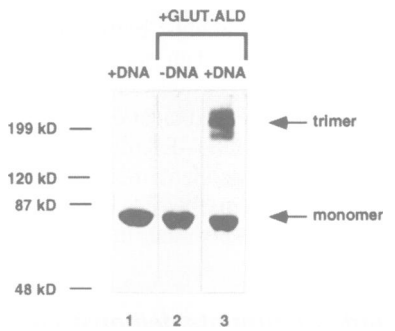


Fig. 4. E1 forms a cross-linked trimer in the presence of DNA. Purified E1 protein was incubated in the absence (lane 2) or presence (lanes 1 and 3) of a DNA fragment containing the ori, followed by crosslinking with glutaraldehyde (lanes 2 and 3). After treatment with Laemmli loading dye, the samples were run on SDS–PAGE followed by a Western blot analysis using a monoclonal antibody directed against E1.

glutaraldehyde. These samples were then run on an SDS–PAGE gel, transferred to nitrocellulose and probed with a monoclonal antibody directed against E1. As demonstrated in Figure 4, a set of bands of ~200 kDa, in addition to monomeric E1, could be detected in the presence of both DNA and crosslinker (Figure 4, lane 3), but not in the absence of DNA (Figure 4, lane 2) or in the absence of crosslinker (Figure 4, lane 1). These results demonstrate that a crosslinked multimer of E1 with a molecular weight consistent with a trimer could form in the presence of ori

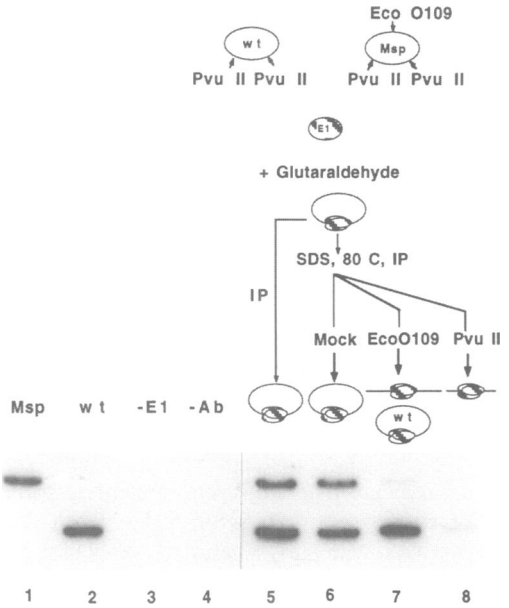


Fig. 5. E1 is linked topologically to plasmid DNA in the E1–ori complex. E1 was bound to the origin of replication in two different plasmids, wt and Msp, which differ in the size of the ori fragment. The Msp plasmid also has a restriction site for *Eco*O109 in the plasmid backbone, while this site has been mutated in the wt ori plasmid. After crosslinking, a fraction of the sample was immunoprecipitated with a monoclonal antibody directed against E1 (lane 5). The rest of the sample was treated with 0.1% SDS and heated at 80°C to denature the protein, followed by immunoprecipitation with the E1 antibody. A fraction of the beads was then either mock digested (lane 6), digested with *Eco*O109 (lane 7) which cleaves the Msp plasmid or with *Pvu*II (lane 8) which cleaves both plasmids, and washed. After treatment with proteinase K and phenol extraction, the presence of DNA was detected by PCR using universal primers. Lanes 1 and 2 are markers for the ori fragment generated by the two input plasmids respectively. Lanes 3 and 4 are controls identical to lane 5 except that either E1 protein or antibody was left out of the reactions.

DNA. Furthermore, consistent with the results from a glycerol gradient analysis, the failure to generate any crosslinked species of E1 larger than the monomer in the absence of DNA indicated that E1 was present exclusively in its monomeric form in the preparation.

The trimeric form of E1 is linked topologically to the ori DNA

To determine if the trimer of E1 forms a ring-like structure encircling the DNA, as predicted from the hydroxyl-radical footprints, we performed the experiment shown schematically in Figure 5 to determine if E1 could be linked topologically to a plasmid containing an E1 binding site. Two different plasmids that contained an ori fragment were incubated with E1 protein under binding conditions and crosslinked with glutaraldehyde. These two pUC 19 plasmids, wt and Msp, differ in the size of their ori fragment; in addition, a site for the restriction enzyme *Eco*O109 in the plasmid backbone of the wt plasmid has been mutated. After crosslinking, the DNA–protein mixture was divided into several samples and treated according to the scheme in Figure 5. One sample was immunoprecipitated with a monoclonal antibody directed against E1 (Figure 5, lane 5). Both plasmids could be immunoprecipitated together with E1, and could subsequently be detected by PCR. In the absence of either E1 or antibody, no DNA was precipitated (Figure 5, lanes 3 and 4). The other samples were heated in SDS to denature the E1 protein. After dilution of the SDS, these samples were immunoprecipitated with the E1 antibody, followed by collection of the immunocomplexes on Protein A–Sepharose beads. These samples were either mock digested or digested with *Eco*O109, which cleaves the Msp plasmid but not the wt plasmid, or with *Pvu*II, which cleaves both plasmids. The digestions were followed by several washes of the beads. As shown in Figure 5, lane 6, both plasmids could be immunoprecipitated and recovered after mock digestion and washes, indicating that E1 was either covalently associated with the DNA or that E1 was crosslinked into a form that is linked topologically to the circular DNA. Digestion with *Eco*O109 (Figure 5, lane 7), which linearizes the Msp plasmid but does not cleave the wt plasmid, resulted in loss of the ability to recover the linearized plasmid but not the circular plasmid. These results indicate that while E1 is firmly associated with circular plasmid DNA, even after treatment with SDS, subsequent linearization of the plasmid leads to a loss of this association. When both plasmids were digested (Figure 5, lane 8) neither could be recovered after the washes, demonstrating that E1 was not linked covalently to the DNA. These results demonstrate that the DNA could not be released from the crosslinked E1 when the DNA was in a circular form, but that the DNA was readily released when the DNA was linearized. The clear implication of these results is that E1 is linked topologically to the DNA in the crosslinked form, and that when the DNA is linearized, E1 can slide off the ends of the DNA molecule. These results are consistent with a model where E1 forms a trimeric ring-like structure that encircles the DNA.

Discussion

The papillomaviruses represent an interesting system for studying the involvement of transcription factors in DNA replication. The absolute requirement for E2 for DNA replication demonstrates that the E2 protein performs an essential function *in vivo*. We have proposed previously that this essential function is to serve as a specificity factor for the binding of E1 to the ori. This conclusion was based on the observation that E1 can bind to the ori with relatively high affinity but shows poor sequence specificity. However, binding of E1 in the presence of E2 resulted in a several hundred-fold increase in sequence specificity (Sedman and Stenlund, 1995). These results are consistent with the notion that an important function for cooperative DNA binding is to generate increased sequence specificity (Ptashne, 1992).

The results we report here provide some explanation for the observed properties of the E1–E2–ori and E1–ori complexes. It has been clear for some time that E2, in addition to the quantitative effect that can be measured as cooperative binding, also has a qualitative effect on DNA binding by E1, as indicated by the reduced apparent molecular weight of the E1–ori complex formed in the presence of E2 (Lusky *et al.*, 1994; Sedman and Stenlund, 1995). As demonstrated here, this difference can be accounted for because E1 binds in a monomeric from in the presence of E2 and in a trimeric form in its absence.

Interestingly, in the absence of E2 we have not, under any conditions, been able to detect any DNA binding activity of E1 in the monomeric form (J.Sedman and A.Stenlund, unpublished data). This observation indicates that E2 serves a critical role in monomeric E1 binding and that the interaction with E2 may induce structural changes in E1, allowing DNA binding.

Several different proteins involved in DNA replication and recombination have been demonstrated to exist as ring-like structures that encircle the double helix. The best studied examples are the polymerase processivity factors proliferating cell nuclear antigen (PCNA) and the β-subunit of *Escherichia coli* polymerase III, which have been termed sliding clamps (Kong *et al.*, 1992; Kuriyan and O'Donnell, 1993; Krishna *et al.*, 1994). These proteins show no apparent sequence specificity in binding to DNA, form 'donuts' in solution and require a specialized machinery to be loaded onto DNA (Stukenberg *et al.*, 1991, 1994; Burgers and Yoder, 1993). The initiator E1 corresponds to a different class of ring-like molecule, and the E1 protein does not form trimers spontaneously in the absence of DNA. Instead, our results indicate that the trimeric E1 ring is generated by the assembly of individual monomers onto DNA, presumably directed by the presence of overlapping binding sites for the monomers of E1. This indicates that the affinity of E1 for DNA in the formation of the E1–ori complex involves a combination of protein–protein interactions and protein–DNA interactions. The ring-like structure may also explain the relatively low sequence specificity of the trimeric E1 complex because it is likely that part of the affinity for DNA is contributed by phosphate backbone contacts and by the protein–protein interactions.

The ring-like structure of the trimeric E1 complex is intriguing. It has been suggested that the hexameric ring-like structures, which can be formed by T-antigen and carry the helicase activity, function by encircling the DNA molecule (Mastrangelo *et al.*, 1989; Borowiec *et al.*, 1990). Other helicases, including the *E.coli* Ruv B branch migration protein, is seen by electron microscopy to form a hexameric structure that encircles the DNA and is likely to function by moving along the DNA (Stasiak *et al.*, 1994). Thus, the formation of a ring-like structure around the DNA may be a common mechanism for loading helicases as well as other factors requiring continuous association with the DNA helix. In contrast to the DNA helicases that have been isolated from *E.coli* which are targeted to DNA by other proteins (Matson and Kaiser-Rogers, 1990), the initiator proteins from DNA tumor viruses, including the E1 protein, specifically recognize the ori. However, the helicase activity of, for example, T-antigen is not strictly dependent on a particular DNA sequence (Dean *et al.*, 1987; Dodson *et al.*, 1987; Scheffner *et al.*, 1989). This change in specificity is not entirely understood in molecular terms, but while, for example, T-antigen can bind DNA as a dimer, the helicase activity is associated with a hexameric form of the protein, as determined by sucrose gradient centrifugation (Wessel *et al.*, 1992). Thus, conversion from a site-specific DNA binding protein to a helicase involves multimerization of the T-antigen. It is likely that *in vivo* this multimerization process takes place on the ori, although high concentrations of T-antigen *in vitro* can form hexameric structures in the absence of the ori DNA.

Based on these results we believe that the highly sequence-specific E1–E2–ori complex serves to recognize the ori and to deposit a single molecule of E1 for high sequence-specific recognition of the origin. In turn, this E1 molecule may serve as a precursor for the formation of the trimeric E1 complex. The trimeric E1 complex, because of its low specificity for the ori, is unlikely to perform this function directly *in vivo*, but the ring-like structure ensures continuous association with the DNA. At present we have no evidence that the trimeric form of E1 can travel along DNA or that it can serve to unwind DNA, and thus the trimeric E1 complex is unlikely to represent the final form of E1 that is active in replication as a DNA helicase. Instead, recent results from our laboratory indicate that a discrete hexameric form of E1 can be generated, and that this form has DNA-dependent ATPase activity which is a hallmark of helicase activity (J.Sedman and A.Stenlund, unpublished data). This type of model is consistent with the results described by Lusky *et al.* (1994), who observed a general correlation between the conditions required for *in vitro* replication and unwinding activity and those required for the generation of larger E1–ori-containing complexes. Based on these observations, Lusky *et al.* (1994) proposed a model where the smaller E1 complex, which also contains E2, serves as a precursor for the formation of a larger E1–ori complex, which has unwinding activity. However, the relationship between those two E1-containing complexes and the complexes we have analyzed here is not entirely clear. The complexes analyzed by Lusky *et al.* (1994) were formed under conditions using much higher protein concentrations and in the presence of magnesium and ATP. In addition, the estimates of stoichiometry that were made by Lusky *et al.* (1994) differ significantly from the measurements obtained here, indicating that these complexes are different. We expect that further biochemical analyses of E1–ori complexes formed under different conditions will establish the composition and activity of these different complexes.

Materials and methods

Molecular mass determination by sedimentation analysis

E1 and E2 proteins were expressed in *E.coli* and purified to apparent homogeneity (J.Sedman and A.Stenlund, submitted). E1–ori and E1–E2–ori complexes were formed in buffer KP [20 mM K-phosphate, pH 7.5, 100 mM NaCl, 1 mM EDTA, 2 mM dithiothreitol (DTT), 0.1% NP-40, 10% glycerol, 100 μg/ml bovine serum albumin (BSA)]. For the E1–ori complex formation, 5 ng E1 protein and 10 pg BPV ori fragment were incubated in a 20 μl volume for 10 min at room temperature in the presence of 20 ng nonspecific competitor DNA. The DNA–protein complexes were stabilized by crosslinking with 2 mM glutaraldehyde for 10 min. Crosslinking was stopped by the addition of Tris–HCl, pH 7.5, to 25 mM. E1–E2–ori complexes were formed under identical conditions but using 1 ng E1 and 10 ng E2 protein. The labeled ori fragment was generated by PCR using labeled universal primers and the ori plasmid wt as template. This plasmid contains a 60-nucleotide BPV ori fragment (nucleotides 7914–27) cloned between the *Xba*I and *Hin*dIII sites in the pUC 19 polylinker.

To determine the sedimentation coefficients of ori–protein complexes, 200 μl of the binding reactions were analyzed on 20–40% glycerol gradients in 20 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA and 2 mM DTT (buffer TS). The gradients were centrifuged at 50 000 r.p.m. for 15 h at 4°C using SW55 Ti rotor, and divided into 0.2 ml fractions for the subsequent analysis. The peak fractions of ori-containing

complexes were determined by measuring radioactivity in each collected fraction and verified by gel electrophoresis. The sedimentation of standard marker proteins was determined using the Bradford assay. The following standard proteins were used: ovalbumin (3.6S), serum albumin (4.3S), aldolase (8.3S), catalase (11.3S) and thyroglobulin (18.5S).

Gel filtration

Sephacryl S400 chromatography was performed to measure the Stoke's radii of the complexes. 200 µl of the binding reactions were loaded onto a 7×300 mm S400 gel filtration column equilibrated with buffer TS. The column was developed with the same buffer. Aliquots of gel filtration chromatography fractions were analysed on 0.8% agarose gels. The elution of standard proteins was determined using the Bradford assay. The following standard proteins were employed: ferritin ($R = 61\times10^{-8}$ cm), catalase ($R = 52.2\times10^{-8}$ cm), BSA ($R = 35.5\times10^{-8}$ cm) and ovalbumin ($R = 30.5\times10^{-8}$ cm). The molecular weights of the complexes were calculated according to the established relationship between Stoke's radius, the sedimentation coefficient and molecular weight (Li and Desiderio, 1993). The partial specific volume of the E1–E2–ori complex was calculated to be 0.67 cm^3/ml, assuming additivity of molar values and a value of 0.73 cm^3/ml for protein and 0.51 cm^3/ml for DNA. The partial specific volume for the E1–ori complex was calculated to be 0.69 cm^3/ml.

Supershift assays

GST–E1 and E1 proteins were mixed in different ratios in the presence of E2, and the E1–E2–ori complexes were formed as described above. The total amount of E1 DNA binding activity was kept constant in all samples. To one-half of each sample, 20 ng anti-GST monoclonal antibody (Santa Cruz Biotechnology) were added 5 min before glutaraldehyde crosslinking treatment of the protein–ori DNA complexes. The resulting complexes were resolved on 0.8% agarose gels in TAE buffer and quantitated by Fuji BAS 1000.

Hydroxyl-radical footprinting

The complex formation reactions were performed as described above, but scaled up to a volume of 600 µl. Complexes were isolated by agarose gel electrophoresis and transferred by capillary blotting onto nitrocellulose. Filters were exposed briefly to film and strips of nitrocellulose containing the complex excised from the filter. Chemical cleavage reactions were carried out as described previously (Dixon *et al.*, 1991). Several small pieces of filter (~3×10 mm) containing the DNA–protein complexes were submerged in 200 µl KP buffer. Cleavage reactions were initiated by the addition of 4 mM sodium ascorbate, 0.8 mM ferrous ammonium sulfate [$(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$], 1.6 mM EDTA and 0.01% hydrogen peroxide. After 2 min at room temperature, the cleavage reaction was terminated with 20 mM thiourea. Modified DNA was then released by soaking the filters in 10 mM Tris, 100 mM NaCl, 1 mM EDTA, pH 7.5, containing 0.1% SDS, and purified by phenol–chloroform extraction and two ethanol precipitations. The cleavage products were analyzed on 8% urea–PAGE gels. G-Specific ladders prepared according to the Maxam–Gilbert sequencing protocol were used as size markers.

Analysis of crosslinked E1 complexes

E1 protein (20 ng) was incubated with a BPV ori fragment in 20 µl buffer containing 20 mM HEPES, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.01% NP-40, 2 mM DTT and 100 µg/ml BSA. Crosslinking with glutaraldehyde was performed by adding glutaraldehyde to a 0.2 mM final concentration. SDS Laemmli loading buffer was added and the samples were heated at 80°C for 5 min. Aliquots were run on 7.5% Laemmli gels and analyzed by Western blotting using a monoclonal E1 antibody. Bio-Rad prestained protein markers were used as molecular weight standards.

Assays for topological linkage between E1 and DNA

The E1–ori complex formation reactions were scaled up to 100 µl using 200 ng each of two circular ori plasmids, wt and Msp. The wt plasmid contained the ori sequences between nucleotides 7914 and 27, and the Msp plasmid contained the ori sequences between nucleotides 7903 and 81, both cloned into the polylinker of pUC 19. SDS denaturation of the E1–ori complexes was performed after the crosslinking reaction by heating at 80°C for 5 min, after the addition of SDS to 0.1%. The sample was diluted 5-fold with KP buffer, and then E1 antibody was added. Complexes were collected with Protein A–Sepharose. The beads were washed three times with KP buffer. Co-precipitated DNA was then purified by proteinase K treatment, phenol extraction and ethanol precipitation. When co-precipitated DNA was subjected to cleavage with restriction nuclease, Protein A–Sepharose beads carrying the E1–protein complex were washed after the binding reaction, as described above, and resuspended in 50 µl of the appropriate buffer. 20 units of enzyme (*Pvu*II or *Eco*O109) were added. The beads were mixed carefully and incubated at 37°C for 15 min, followed by three washes to remove DNA released by the restriction nuclease treatment. Of the recovered sample, 2% was used as template for 10 amplification cycles of PCR with labeled primers.

Acknowledgements

We thank B.Stillman and M.Tanaka for critically reading this manuscript. This work was supported by a grant to A.S. from the National Institutes of Health, CA 13106.

References

Androphy,E.J., Lowy,D.R. and Schiller,J.T. (1987) Bovine papillomavirus E2 *trans*-activating gene product binds to specific sites in papillomavirus DNA. *Nature*, **325**, 70–73.

Blitz,I. and Laiminis,L. (1991) The 68-kilodalton E1 protein of bovine papillomavirus is a DNA binding phosphoprotein which associates with the E2 transcriptional activator *in vitro*. *J. Virol.*, **65**, 649–656.

Borowiec,J.A., Dean,F.B., Bullock,P.A. and Hurwitz,J. (1990) Binding and unwinding — How T-antigen engages the SV40 origin of DNA replication. *Cell*, **60**, 181–184.

Burgers,P.M.J. and Yoder,B.L. (1993) ATP-independent loading of the proliferating cell nuclear antigen requires DNA ends. *J. Biol. Chem.*, **268**, 19923–19926.

Dean,F., Bullock,P., Murakami,Y., Wobbe,C.R., Weissbach,L. and Hurwitz,J. (1987) Simian virus 40 (SV40) DNA replication: SV40 large T antigen unwinds DNA containing the SV40 origin of replication. *Proc. Natl Acad. Sci. USA*, **84**, 16–20.

Dixon,W.J., Hayes,J.J., Levin,J.R., Weidner,M.F., Dombroski,B.A. and Tullius,T.D. (1991) Hydroxyl radical footprinting. *Methods Enzymol.*, **208**, 380–413.

Dodson,M., Dean,F., Bullock,P., Echols,H. and Hurwitz,J. (1987) Unwinding of duplex DNA from the SV40 origin of replication by T antigen. *Science*, **238**, 964–967.

Fanning,E. and Knippers,R. (1992) Structure and function of simian virus 40 large tumor antigen. *Annu. Rev. Biochem.*, **61**, 55–85.

Gillette,T.G., Lusky,M. and Borowiec,J.A. (1994) Induction of structural changes in the bovine papillomavirus type 1 origin of replication by the viral E1 and E2 proteins. *Proc. Natl Acad. Sci. USA*, **91**, 8846–8850.

Holt,S.E. and Wilson,V.G. (1995) Mutational analysis of the 18-basepair inverted repeat element at the bovine papillomavirus origin of replication: identification of critical sequences for E1 binding and *in vivo* replication. *J. Virol.*, **69**, 6525–6532.

Holt,S.E., Schuller,G. and Wilson,V.G. (1993) DNA binding specificity of the bovine papillomavirus E1 protein is determined by the sequences contained within an 18-base-pair inverted repeat element at the origin of replication. *J. Virol.*, **69**, 6525–6532.

Kong,X.-P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) Three-dimensional structure of the β-subunit of *E. coli* DNA polymerase III holoenzyme: a sliding clamp. *Cell*, **69**, 425–437.

Krishna,T.S.R., Kong,X.-P., Gary,S., Burgers,P. and Kuriyan,J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*, **79**, 1233–1243.

Kuriyan,J. and O'Donnell,M. (1993) Sliding clamps of DNA polymerases. *J. Mol. Biol.*, **234**, 915–925.

Le Moal,M.A., Yaniv,M. and Thierry,F. (1994) The bovine papillomavirus type 1 (BPV1) replication protein E1 modulates transcriptional activation by interacting with BPV 1 E2. *J. Virol.*, **68**, 1085–1093.

Li,M. and Desiderio,S. (1993) Physical characterization of DNA binding proteins in crude preparations. In Harris,E.L.V. and Angal,S. (eds), *Transcription Factors: A Practical Approach*. IRL Press at Oxford University Press, Oxford, UK, pp. 185–196.

Lusky,M. and Fontane,E. (1991) Formation of the complex of bovine papillomavirus E1 and E2 proteins is modulated by E2 phosphorylation and depends on sequences within the carboxyl terminus of E1. *Proc. Natl Acad. Sci. USA*, **88**, 6363–6367.

Lusky,M., Hurwitz,J. and Seo,Y.-S. (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.

Mastrangelo,I., Hough,P., Wall,J., Dodson,M., Dean,F. and Hurwitz,J. (1989) ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature*, **338**, 658–662.

Matson,S.W. and Kaiser-Rogers,K.A. (1990) DNA helicases. *Annu. Rev. Biochem.*, **59**, 289–329.

McBride,A.A., Romanczuk,H.R. and Howley,P.M. (1991) The papillomavirus E2 regulatory proteins. *J. Biol. Chem.*, **266**, 18411–18414.

Mohr,I.J., Clark,R., Sun,S., Androphy,E.J., MacPherson,P. and Botchan,M. (1990) Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator. *Science*, **250**, 1694–1699.

Ptashne,M. (1992) *A Genetic Switch*. Cell Press, Cambridge, MA.

Scheffner,M., Wessel,R. and Stahl,H. (1989) Sequence independent duplex DNA opening reaction catalyzed by SV40 large tumor antigen. *Nucleic Acids Res.*, **17**, 93–106.

Sedman,J. and Stenlund,A. (1995) Cooperative interaction between the initiator E1 and the transcriptional activator E2 is required for replication of bovine papillomavirus *in vivo* and *in vitro*. *EMBO J.*, **14**, 6218–6228.

Seo,Y.-S., Muller,F., Lusky,M. and Hurwitz,J. (1993a) Bovine papillomavirus (BPV)-encoded E1 protein contains multiple activities required for BPV DNA replication. *Proc. Natl Acad. Sci. USA*, **90**, 702–706.

Seo,Y.-S., Muller,F., Lusky,M., Gibbs,E., Kim,H.-Y., Phillips,B. and Hurwitz,J. (1993b) Bovine papillomavirus (BPV)-encoded E2 protein enhances binding of E1 protein to the BPV replication origin. *Proc. Natl Acad. Sci. USA*, **90**, 2865–2869.

Spalholz,B.A., Yang,Y.C. and Howley,P.M. (1985) Transactivation of a bovine papillomavirus transcriptional regulatory element by the E2 gene product. *Cell*, **42**, 183–191.

Stasiak,A., Tsaneva,I.R., West,S.C., Benson,C.J.B., Yu,H. and Egelman,H. (1994) The *Escherichia coli* RuvB branch migration protein forms double hexameric rings around DNA. *Proc. Natl Acad. Sci. USA*, **91**, 7618–7622.

Stukenberg,P.T., Studwell-Vaughan,P.S. and O'Donnell,M. (1991) Mechanism of sliding β-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.*, **266**, 11328–11334.

Stukenberg,P.T., Turner,J. and O'Donnell,M. (1994) An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. *Cell*, **78**, 877–887.

Thorner,L., Lim,D. and Botchan,M. (1993) DNA binding domain of bovine papillomavirus type 1 E1 helicase: structural and functional aspects. *J. Virol.*, **67**, 6000–6014.

Ustav,M. and Stenlund,A. (1991) Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *EMBO J.*, **10**, 449–457.

Ustav,M., Ustav,E., Szymanski,P. and Stenlund,A. (1991) Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J.*, **10**, 4321–4329.

Ustav,E., Ustav,M., Szymanski,P. and Stenlund,A. (1993) The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. *Proc. Natl Acad. Sci. USA*, **90**, 898–902.

Wessel,R., Schweizer,J. and Stahl,H. (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.

Wilson,V.G. and Ludes-Meyers,J. (1991) A bovine papillomavirus E1-related protein binds specifically to bovine papillomavirus DNA. *J. Virol.*, **65**, 5314–5322.

Yang,L., Li,R., Mohr,I., Clark,R. and Botchan,M.R. (1991) Activation of BPV-1 replication *in vitro* by the transcription factor E2. *Nature*, **353**, 628–633.

Yang,L., Mohr,I., Fouts,E., Lim,D.A., Nohaile,M. and Botchan,M. (1993) The E1 protein of the papillomavirus BPV-1 is an ATP-dependent DNA helicase. *Proc. Natl Acad. Sci. USA*, **90**, 5086–5090.

Received on May 2, 1996; revised on June 12, 1996