The EMBO Journal vol.14 no.24 pp.6218–6228, 1995

Co-operative interaction between the initiator E1 and the transcriptional activator E2 is required for replicator specific DNA replication of bovine papillomavirus *in vivo* and *in vitro*

Juhan Sedman and Arne Stenlund[1]

Cold Spring Harbor Laboratory, 1 Bungtown Road, PO Box 100, Cold Spring Harbor, NY 11724, USA

[1]Corresponding author

The E1 polypeptide from bovine papillomavirus binds to the origin of replication (ori) and possesses the activities attributed to initiator proteins. E1 is also the only viral protein required for replication in a cell-free replication system. Replication *in vivo*, however, absolutely requires in addition the viral transcription factor E2. We demonstrate that the basis for this distinction between *in vitro* and *in vivo* requirements is the limited sequence specificity of the E1 protein. E1 and E2, which bind the ori individually with low sequence specificity, together bind with greatly increased sequence specificity. This combinatorial effect provides a function for the involvement of transcriptional activation domains in replication and suggests common mechanisms of action for transcription factors in both transcription and replication. It also provides a possible explanation for the differential specificity that is observed for auxiliary transcription factors *in vivo*.

Keywords: DNA replication/papillomavirus/transcription factor

Introduction

A general feature of hitherto characterized eukaryotic replicons is a link between replication and transcription such that transcription factor binding sites form a part of the *cis*-acting elements (ori) required for initiation of replication *in vivo* (for a recent review, see DePamphilis, 1993). The ori, for full activity, depends on transcription factors bound to these sites but the degree of dependence on these transcription factors varies depending on the replicon. Recent work with some viral replicons has demonstrated that, even though in some cases several different transcription factors can be used interchangeably, in other cases only a few factors are functional (Cheng and Kelly, 1989; Baru *et al.*, 1991; Bennet-Cook and Hassel, 1991; Guo and DePamphilis, 1992). In addition, some transcription factors that function for one replicon lack activity for another replicon, demonstrating some degree of specificity (Guo and DePamphilis, 1992).

In the few cases where cell-free replication systems have been developed, little or no dependence on the transcription factor is evident. One property that has been suggested to account for this difference is that *in vivo* the template DNA exists in a nucleosome-bound form while in cell-free systems naked DNA is generally used. It has been proposed that transcription factors bound to the ori could serve to prevent repression by nucleosomes (Cheng and Kelly, 1989; Cheng *et al.*, 1992; Li and Botchan, 1994). However, this model does not explain the observed specificity of some transcription factors for certain oris.

In bovine papillomavirus the dependence on a transcription factor *in vivo* appears to be absolute; no replication can be detected in the absence of E2 (Ustav and Stenlund, 1991). This requirement is also highly specific. The required transcription factor, E2, is encoded by the virus (reviewed in McBride *et al.*, 1991) and other transcription factors that have been tested are inactive. For example, a hybrid protein where the activation domain of E2 has been replaced with the activation domain from the herpes virus protein VP16 is inactive for replication, demonstrating the importance of the E2 activation domain (Ustav *et al.*, 1991). Furthermore, the polyomavirus enhancer can not functionally replace the E2 binding site in spite of its ability to bind a variety of cellular transcription factors in rodent cells (M.Ustav and A.Stenlund, unpublished results). In contrast to the requirement for E2 for replication *in vivo*, very modest effects of added E2 are generally observed for replication *in vitro* and dependence on the presence of an E2 binding site at the origin of replication has not been demonstrated (Yang *et al.*, 1991; Seo *et al.*, 1993a; Lusky *et al.*, 1994). This is in spite of the fact that co-operativity in binding and also DNA-independent interaction between E1 and E2 can be observed *in vitro* (Mohr *et al.*, 1990; Lusky and Fontane, 1991; Yang *et al.*, 1991; Seo *et al.*, 1993b; Gillette *et al.*, 1994; Le Moal *et al.*, 1994). The ~20-fold co-operativity that can be observed for binding *in vitro* does not adequately explain the absolute dependence on E2 *in vivo*, since overexpression of E1 to very high levels does not alleviate the requirement for E2 (M.Ustav and A.Stenlund, unpublished results). Thus, although E1 *in vitro* in the absence of E2 has all the activities normally associated with initiators (Seo *et al.*, 1993a; Yang *et al.*, 1993; Park *et al.*, 1994), this is insufficient for replication *in vivo*.

Here, we resolve this issue and present evidence that under appropriate conditions, a specific multiprotein complex containing both E1 and E2 forms on the ori in the presence of binding sites for the two proteins. The ability to form this particular complex shows a very strong correlation with replication *in vivo* and is dependent on specific interactions between the two proteins and their recognition sequences in the DNA, as well as interactions between the two proteins. One of the striking consequences of these interactions is a substantial increase in the sequence specificity of the complex compared with that of the individual components. Under conditions where ori binding is limited by the sequence specificity of E1, replication *in vitro* can be made completely dependent on the presence of the E2 transcription factor and an

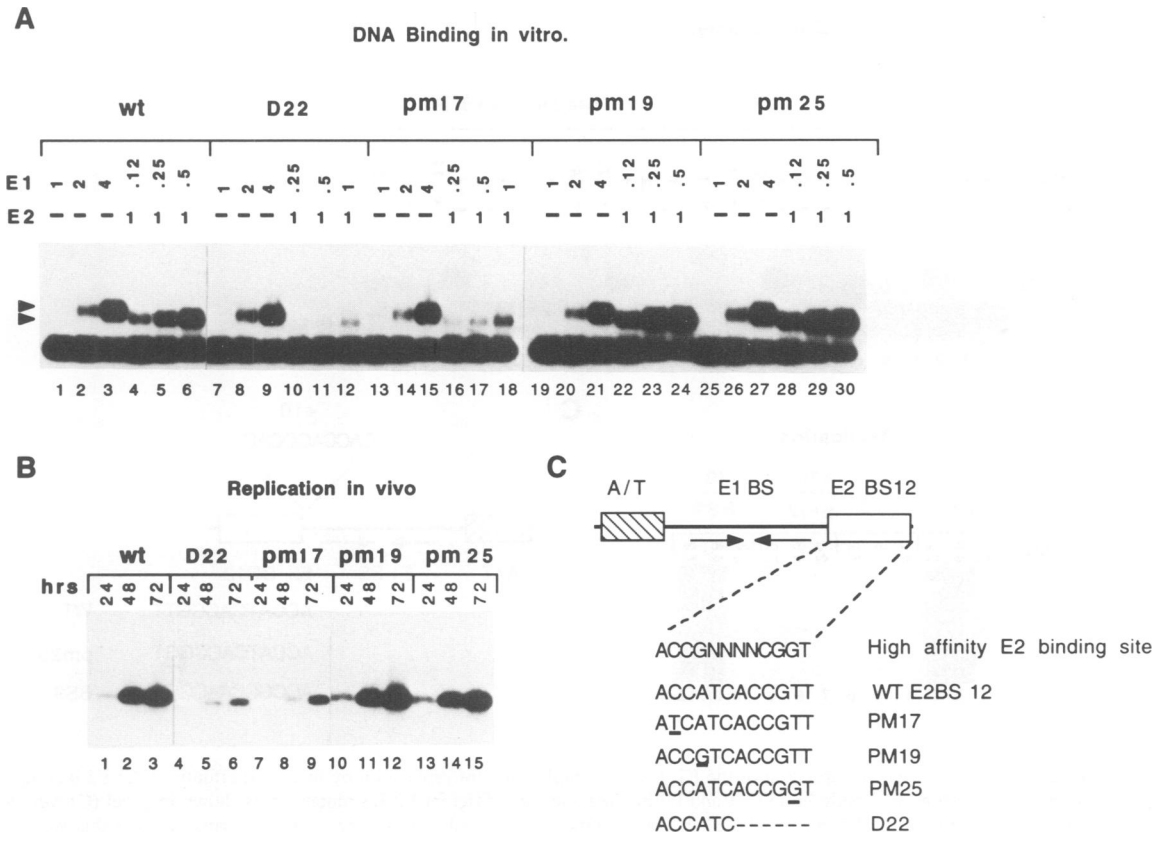

Fig. 1. DNA replication activity *in vivo* correlates with the ability to form an E1–E2–ori complex *in vitro*. (**A**) Oris with mutations in the E2 binding site as depicted schematically in panel (C) were tested for their ability to direct the formation of the E1–ori and the E1–E2–ori complexes in a gel-shift assay. The purified E1 and E2 proteins were incubated with ori probes followed by cross-linking with glutaraldehyde and the resulting complexes were separated on agarose gels. At high concentrations E1 alone can bind to the ori probe to form the E1–ori complex. The titration points were chosen such that at the lowest concentration of E1 no complex is formed and this lane serves as a probe alone control. In the presence of E2 another complex (E1–E2–ori) can form at lower E1 concentrations. This latter complex has higher mobility than the E1–ori complex. The two arrowheads indicate the difference in migration between the E1–ori and E1–E2–ori complexes. (**B**) The same mutants as in panel (A) were tested for replication in the transient replication assay by transfection of 100 ng of the ori constructs into the 4.15 cell line that constitutively expresses the BPV E1 and E2 proteins. Time points were taken at 48 and 72 h after transfection, and replicated, *Dpn*I-resistant DNA was analyzed by Southern analysis. (**C**) A schematic drawing of the minimal ori is shown illustrating the position of the A/T-rich region, the E1 binding site, and the E2 binding site. The sequences of the mutants in the E2 BS are shown, together with a consensus high-affinity binding site. The two mutants 17 and 22 have reduced affinity for E2, the two mutants 19 and 25 have increased affinity for E2 compared with the wt E2 BS12.

appropriately positioned binding site for E2. We propose that this may be a common function for transcription factors also in other replicons and suggest that the strategy of using two or more DNA binding factors with adjacent sites that can interact with each other, may be a universal strategy to achieve a high degree of sequence-specific DNA binding also in other regulatory systems.

Results

Replication in vivo correlates with the ability to form a specific E1–E2–ori complex in vitro

Studies of the DNA-binding activity of the E1 protein have indicated that the E1 protein can bind to an E1 binding site in the minimal ori by itself (Ustav *et al.*, 1991; Wilson and Ludes-Meyers, 1991; Yang *et al.*, 1991; Holt *et al.*, 1994), or E1 and E2 can bind together to the adjacent E1 and E2 binding sites that are present in the minimal ori. Binding of E1 alone or E1 and E2 results in two different complexes with different characteristics (Yang *et al.*, 1991; Spalholz *et al.*, 1993; Lusky *et al.*, 1994; J.Sedman and A.Stenlund, unpublished results). These two complexes can be distinguished by their different mobilities in gel-shift assays. The complex formed by E1 alone (E1–ori) migrates more slowly than the complex formed by E1 and E2 in combination (E1–E2–ori), indicating that E1 binds with different stoichiometry in these two complexes. To establish a correlation between the ability to form these two complexes *in vitro* and replication *in vivo* we adopted a genetic approach. We tested ori mutants both for their ability to form the two complexes *in vitro* in a gel-shift assay and for their ability to replicate *in vivo* in a transient replication assay.

Four different mutants that all affect the E2 binding site were first tested (Figure 1C). Two of these mutants, a 6 nt deletion, D22, and the point mutation pm17 have a negative effect on E2 binding. Two other mutants, pm19 and pm25 are point mutations that increase the affinity for E2. The ability to form the two complexes was measured by incubating the probes with the indicated purified proteins followed by cross-linking with glutaraldehyde. The cross-linking is required since the complexes appear to be unstable in this gel system. The complexes were then resolved on 1% agarose gels. In the

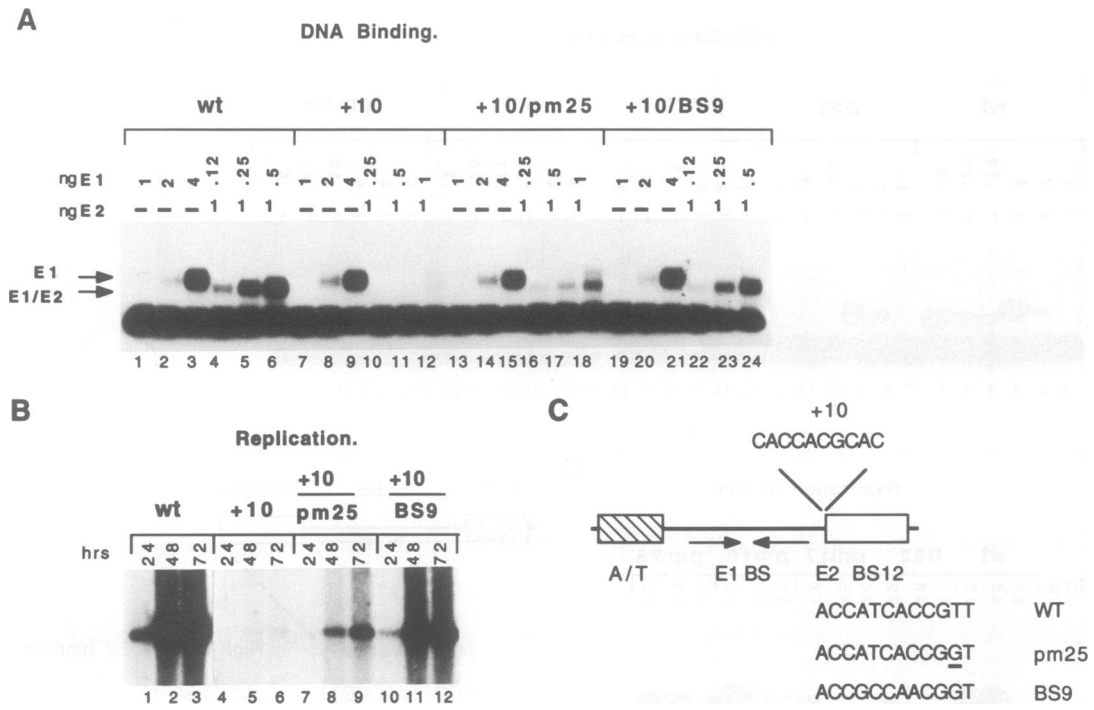

Fig. 2. Spacing mutants can be rescued for formation of the E1–E2–ori complex and for replication by increased affinity of the E2 binding site. (**A**) Origin probes with the insertion +10 between the E1 and E2 binding sites and different E2 BS mutations as shown in panel (C) were tested for their ability to direct the formation of the E1–ori and the E1–E2–ori complexes as described in Figure 1. (**B**) The same mutants that were used in panel (A) were tested for their ability to replicate in a transient replication assay as described in Figure 1. (**C**) The mutations in the ori are illustrated. The inserted sequence was designed to conserve three nucleotides of the flanking sequences of both the E1 BS and the E2 BS.

presence of E1 alone, these probes all form the more slowly migrating E1–ori complex at the same E1 concentration (compare the first three lanes of each set in Figure 1A). However, their ability to form the faster migrating E1–E2–ori complex varied considerably. The wild-type (wt) ori readily forms this complex at a concentration of E1 ~20-fold lower than what is required for formation of the E1–ori complex. (E2 in all cases is present in excess and the amount of complex formed is limited by the quantity of E1 added.) The two E2 BS mutations D22 and pm17 that have reduced affinity for E2 were also reduced in their ability to form the E1–E2–ori complex (15% of wt, lanes 10–12 and 16–18). The two mutations pm19 and pm25 that have an increased affinity for E2 both have a slightly increased ability to form the E1–E2–ori complex (lanes 22–24 and 28–30). The same mutant oris were then tested for replication in a transient replication assay. As shown in Figure 1B, the two mutants with reduced ability to form the E1–E2–ori complex replicate poorly (10% of wt, lanes 4–6 and 7–9), while the two mutants with slightly increased ability to form the E1–E2–ori complex replicate at wt levels (lanes 10–16).

A mutant with increased distance between the E1 and E2 binding sites was then tested as shown in Figure 2. The +10 mutation increases the spacing between the E1 and E2 binding sites by 10 nucleotides. Formation of the E1–ori complex was unaffected by this insertion. However, in spite of the fact that the actual sequence of the E2 binding site is unchanged by the insertion, this ori construct fails to form the E1–E2–ori complex (Figure 2A, lanes 7–12). Replication of this mutant is reduced 100-fold compared with the wt ori (Figure 2B, lanes 4–6). Introduction of a single point mutation at position 25 that increases the affinity for E2 (pm25) restores partially both replication activity (10% of wt; Figure 2B, lanes 7–9) and the ability to form the E1–E2–ori complex (15% of wt; Figure 2A, lanes 16–18). A further increase of the affinity of the E2 binding site through replacement with the high-affinity E2 binding site E2 BS9 resulted in further increased replication (25–30% of wt; Figure 2B, lanes 10–12), as well as increased ability to form the E1–E2–ori complex (50% of wt; Figure 2A, lanes 22–24). Thus, a relationship exists between the affinity of the E2 binding site required for replication and E1–E2–ori complex formation and its distance from the E1 binding site. The explanation for this curious property appears to be that what is important for replication is not binding of E2 *per se*—which is independent of the distance from the E1 binding site—but the ability to form the E1–E2–ori complex, which appears to depend on the distance between the two sites. Thus, replication activity *in vivo* correlates very well with the ability to form the E1–E2–ori complex *in vitro*, which is determined by both the affinity of the E2 binding site and its distance from the E1 binding site.

The ability to form the E2 independent E1 complex on the ori is not required for replication activity in vivo

None of the mutants tested above affected the binding of E1 alone, so the importance of the E1–ori complex for replication could not be determined. We therefore designed mutants that would be capable of forming the E1–E2–ori

complex, but would be defective for E1–ori complex formation. Based on interference analysis (which will be reported elsewhere), the recognition sequence for E1 in both complexes is very similar, indicating that mutations in the E1 BS would affect the formation of both complexes. We reasoned that since the ability to form the E1–E2–ori complex appeared to be a function of the combined binding activities of E1 and E2, it might be possible to rescue a mutant in the E1 binding site, which would affect both the E1–ori and E1–E2–ori complex formation, by increasing the affinity with which E2 binds to the E2 BS. Consequently, we chose two different point mutations (pm8 and pm9) in the E1 BS that are defective for replication, E1–ori complex formation, and E1–E2–ori complex formation and replaced in this context the naturally occurring E2 BS12 with a higher affinity site, E2 BS9.

As can be seen in Figure 3A, both pm8 and pm9 have a reduced ability to form both the E1–ori complex and the E1–E2–ori complex, with pm9 being more severely affected than pm8 (compare lanes 1–6 with lanes 7–12 and 12–18). As expected, the replacement with the high-affinity E2 BS9 did not substantially affect the ability of either of these templates to form the E1–ori complex (Figure 3A, lanes 19–21, 25–27 and 31–33). However, the ability to form the E1–E2–ori complex was significantly improved for both mutants approaching wt levels of complex formation (compare lanes 10–12 with 28–30, and lanes 16–18 with 34–36). When these mutants were tested for replication in the transient replication assay (Figure 3B), the pm 8 mutant was reduced ~10-fold (lanes 5 and 6), and the pm9 mutant was reduced >20-fold for replication (lanes 9 and 10). The replacement with the high-affinity E2 BS in both cases restored replication to within 2-fold of wt (compare lanes 7, 8, 11, 12 with lanes 1 and 2). These results demonstrated very clearly that the ability to form the E1–ori complex did not appear to affect replication, while the ability to form the E1–E2–ori complex did. One conclusion from these results is that the individual binding activities of E1 and E2 are not the final determinant for replication: a weak E1 binding site can be compensated for by a strong E2 binding site, indicating that the E1–E2–ori complex results from the sum of at least three interactions, E1 binding DNA, E2 binding DNA, and an interaction between E1 and E2.

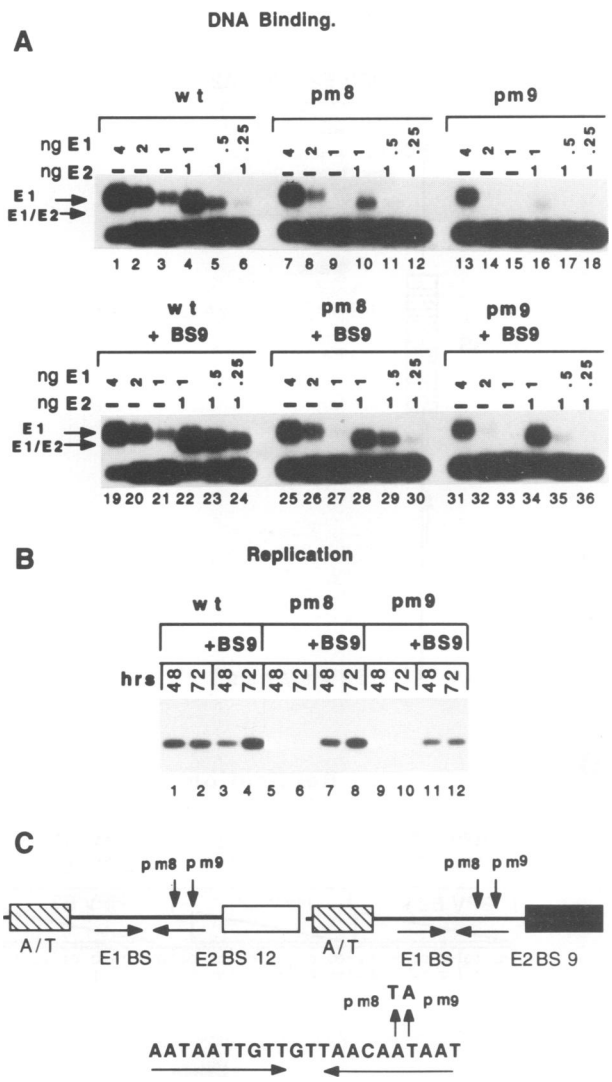

Fig. 3. A high-affinity E2 binding site can restore replication and the ability to form the E1–E2–ori complex for mutants in the E1 binding site. (**A**) Gel-shift assays as described in Figure 1 were performed on origin probes containing either the wt E1 BS, or, either of two point mutations (pm8 or pm9) in the E1 binding site in the context of the wt E2 BS12. In parallel, the wt and the same two mutants were combined with the high-affinity E2 BS9, and tested for their ability to form the E1–ori and E1–E2–ori complexes. (**B**) The same mutants assayed for binding in panel (A) were tested for replication in a transient replication assay as described in Figure 1. (**C**) Schematic illustration of the sequence of the E1 binding site and the location of the point mutations pm8 and pm9.

The related HPV-11 E2 protein fails to co-operate with BPV E1 for binding and does not support BPV ori replication

To test directly the importance of co-operative binding as opposed to mere occupancy of the E2 binding site, we tested the ability of a related E2 protein, HPV-11 E2, to bind co-operatively with BPV E1 on the BPV ori. To ensure that HPV-11 E2 could bind with high affinity, we replaced the E2 BS12 binding site with a high-affinity E2 binding site from HPV-11; this site is also a high-affinity site for BPV E2. As demonstrated in the DNaseI footprints in Figure 4A, both BPV and HPV-11 E2 proteins bind very well to this E2 binding site in the absence of E1 (compare lane 1 with lanes 2 and 6). Upon addition of BPV E1, a combined E1–E2 footprint is generated when BPV E2 is present (lanes 3–5). The HPV-11 E2 protein, however, in spite of efficient binding to the E2 binding site, failed to show any detectable sign of co-operativity with BPV E1 (lanes 7–9), indicating that the two proteins are not interacting. At the highest concentration of E1 (lane 9), E1 alone is capable of weakly protecting the E1 binding site.

In Figure 4B, replication results are shown using the same ori construct that was used for the footprint analysis in Figure 4A. An expression vector for BPV E2 is able to support replication efficiently from the BPV ori (lanes 1–6); HPV E2 likewise supports replication from the HPV ori in the presence of HPV E1, providing a control for expression of functional HPV E2 protein (lanes 7–12). HPV E2, however, fails to support replication from the BPV ori (lanes 13–18). These results demonstrate that binding of E2 to the ori *per se* is not sufficient to activate replication consistent with the results from the spacing

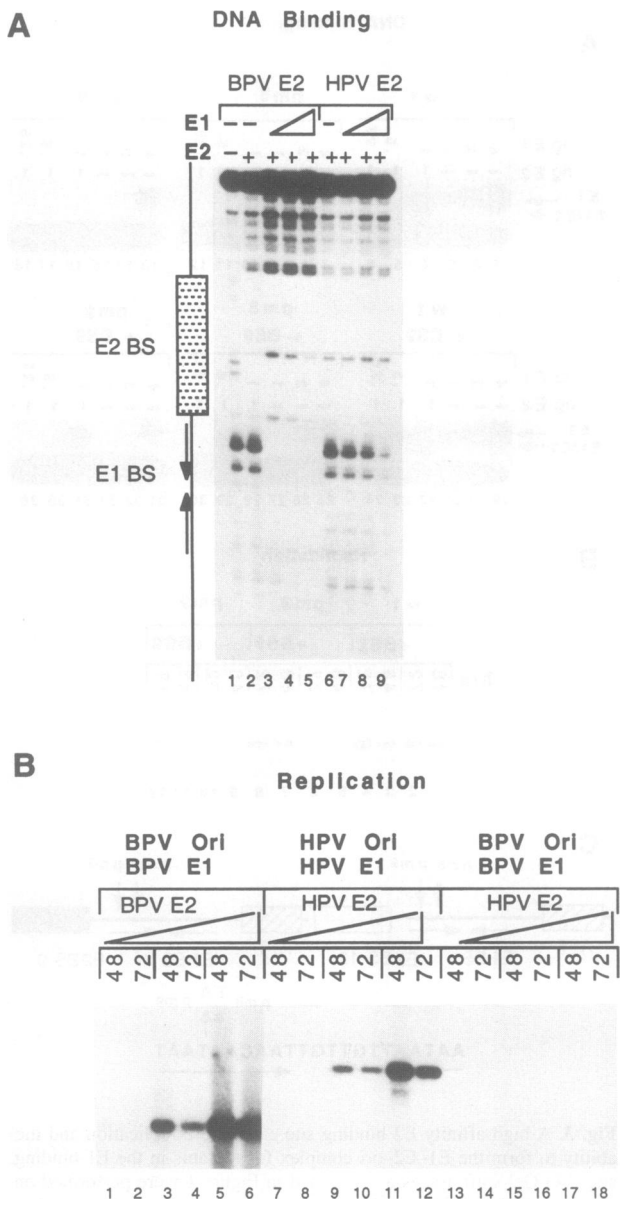

Fig. 4. The E2 protein from HPV-11 fails to form the E1–E2–ori complex and does not function for replication. (**A**) DNase footprint analysis was carried out on an ori fragment where the wt E2 BS12 had been replaced with a high-affinity E2 binding site from the HPV-11 genome. BPV E1 protein was titrated in the absence or the presence of either BPV E2 or HPV-11 E2 proteins as indicated in the figure. E1 was used at 0.5, 1 and 2 ng. Both BPV E2 and HPV E2 were added at 1 ng. The positions of the E1 and E2 binding sites are indicated. (**B**) Transient replication assays were carried out by transfecting CHO cells with various combinations of ori constructs and expression vectors for E1 and E2 from BPV and from HPV-11 as indicated in the figure, varying the quantity of E2 expression vector. Low-molecular weight DNA was harvested at 48 and 72 h after transfection and analyzed for replicating DNA by southern analysis.

mutants and indicating that a specific interaction between E1 and E2 is required for replication.

The combined E1–E2 complex is several hundred fold more sequence specific than either E1 or E2 alone

The apparent requirement for the E1–E2–ori complex for replication *in vivo* presented a paradox of sorts: in a cell-free replication system that recently has been developed, E1 is clearly sufficient for ori-specific replication (Yang *et al.*, 1991, 1993; Seo *et al.*, 1993a; Muller *et al.*, 1994). Since E2 is not required in this system, the E1–E2–ori complex is clearly not required. Hence, replication can be initiated *in vitro* without the formation of the E1–E2–ori complex. A possible resolution to this paradox could be that E2 serves an auxiliary role in replication, and that this requirement can be bypassed *in vitro*. One possibility that occurred to us was that since E1 appears to have a relatively modest sequence specificity (i.e. the ability to distinguish between specific and non-specific sequences), E2 might function to increase the selectivity of binding of E1. The presence of vast quantities of competing cellular DNA sequences under *in vivo* but not *in vitro* conditions could then explain why E2 is required for replication *in vivo* but not *in vitro*.

To test this hypothesis we measured the ability of E1 to bind selectively to the ori, either alone or in combination with E2, in the presence of competitor DNA. This was achieved by measuring binding of E1 in the presence of increasing amounts of a non-specific competitor, poly dI/dC, in the absence or presence of E2. As demonstrated in Figure 5, at this level of E1, in the absence of E2 and in the absence of competitor, E1 protects a large region, including most of the E2 binding site, the E1 binding site, the A+T-rich region and part of the polylinker sequence (compare lanes 1 and 2). In the presence of increasing quantities of competitor DNA (lanes 3–7) the E1 footprint persists only to competitor quantities in the order of 4–8 ng (lanes 4 and 5) which corresponds roughly to a 40- to 80-fold excess of competitor.

E2 alone gives rise to a very weak protection over the low-affinity E2 binding site (lane 8) and binding to this site is also very sensitive to competition. The E2 footprint can be competed by 5 ng of dI/dC (data not shown). In the presence of both E1 and E2 (lanes 10–15), a combined E1+E2 footprint can form. This footprint shows a greatly increased resistance to competitor and persists up to at least 1000 ng of dI/dC, 125–250 times higher competitor concentrations than for E1 alone. Thus, the two proteins that individually bind to these two sites with a low degree of selectivity, in combination bind with several hundred fold greater sequence specificity. Similar experiments were also performed using poly dA/dT and plasmid (pUC 19) as competitors. The results were virtually identical, indicating that this effect is general. However, poly dA/dT is an ~10-fold more potent competitor than dI/dC at the same concentration (data not shown).

Cell-free BPV replication becomes dependent on the E2 transcription factor and the E2 binding site when ori recognition by E1 is challenged

On the basis of these binding studies it was logical to test if selectivity of binding would also be related to the lack of a requirement for E2 in an *in vitro* replication system. We therefore adopted the *in vitro* replication system developed by Yang *et al.* using E1 and E2 proteins expressed and purified from *Escherichia coli*. In the *in vitro* reactions we used three different templates, as shown in Figure 6B. In addition to the wt ori which has a low-affinity site for E2, we also used a mutant where the entire E2 BS was deleted (–E2 BS) as well as an ori

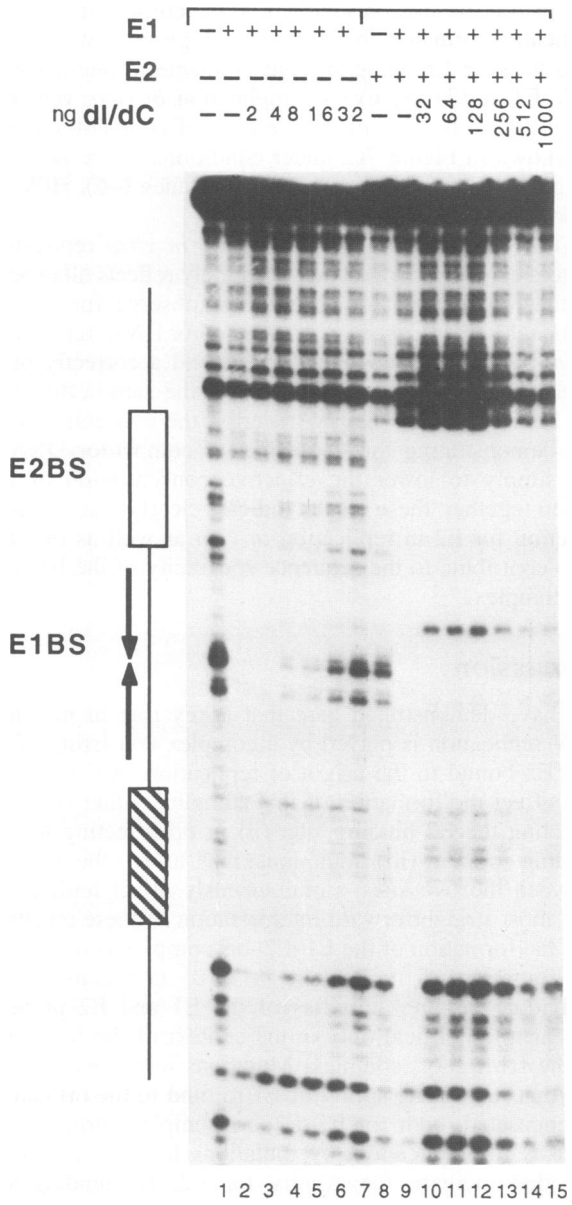

Fig. 5. The E1 and E2 proteins in combination bind the ori with greatly increased sequence specificity. DNase footprint analysis was performed on a minimal origin probe by mixing the probe with increasing quantities of non-specific competitor (poly dI/dC) as indicated. A fixed quantity of either E1 alone (4 ng), or fixed quantities of E1(4 ng) and E2 (1 ng) were then added and probe occupancy was measured by DNase cleavage. The position of the E2 binding site, the E1 binding site and the A+T-rich region are indicated. The extended protection in lanes 14 and 15 is caused by inhibition of DNase cleavage at high competitor concentrations. Note that the footprinting in the presence of both E1 and E2 was performed at substantially higher concentrations of dI/dC than in the presence of E1 alone.

where the low-affinity E2 BS was replaced with a high-affinity site for E2 (wt BS9). As shown in Figure 6A, under these standard conditions in the absence of competitor DNA (–Comp) replication takes place in the absence of E2, and the addition of E2 has a very small effect on replication (compare the first 8 lanes for each template). This slight effect is apparent whether or not an E2 binding site is present on the template and therefore does not appear to be caused by co-operative binding (lanes 17–19 and 21–23).

In the presence of non-specific competitor DNA (+Comp), replication in the absence of E2 is suppressed 50- to 60-fold for all three templates (lanes 9–11, 25–27 and 41–43). Under these conditions, addition of E2 can restore replication to levels similar to those observed without competitor DNA (compare lanes 1–3 with lanes 13–15 and lanes 33–35 with lanes 45–47). This effect of E2 shows dependence on the presence of an E2 binding site on the template. Replication of the wt and wt E2 BS9 templates was stimulated 30- and 45-fold respectively by the addition of E2. The E2 BS deletion showed only a low degree of stimulation (~6-fold). Thus, the presence of E2 made replication *in vitro* insensitive to high concentrations of competitor DNA in a manner that is dependent on the presence of an E2 binding site. These results are consistent with the effects on sequence specificity seen in the DNase footprint assay (Figure 5). Thus, the dependence on E2 and an E2 binding site observed for replication *in vivo* can be reproduced *in vitro* by limiting the ability of E1 to recognize the viral origin of replication through the addition of competitor DNA.

One of the important aspects of the requirement for E2 *in vivo* is the ability of E1 and E2 to interact, rather than just the presence of bound E2 transcription factor at the ori as indicated by the experiments with HPV E2 (Figure 4). To determine how well the more stringent *in vitro* system reflected this aspect of the *in vivo* requirement for E2, we tested several ori mutants and variants that have phenotypes in the *in vitro* binding assays and for replication *in vivo*. The *in vivo* experiments have shown that the E2 binding site functions less well when the distance from the E1 binding site is increased, and that the activity can be restored by increasing the affinity of the E2 binding site (Figure 2). *In vitro*, in the absence of competitor DNA, replication of a template with a moved E2 binding site (+10) is indistinguishable from the wt template both in the absence and presence of E2 even when the high-affinity E2 BS9 is present (Figure 7A, compare lanes 1–6, 13–18 and 25–30). Addition of competitor in the absence of E2 suppresses replication ~40-fold for all templates. Addition of E2 can rescue replication of all three templates to some extent; the replication of the wt template is rescued to 90% of the levels observed without competitor (compare lanes 1–3 with 10–12). The +10 mutant is rescued ~3-fold less efficiently (30% of levels without competitor, compare lanes 13–15 with 22–24). The replacement with a high-affinity E2 BS, +10/BS9, restores replication activity to the spacing mutant which is rescued as effectively as the wt.

We also tested a point mutation in the E1 binding site (pm8) that in the context of the wt E2 BS, is severely reduced in its ability to form the E1–E2–ori complex *in vitro* and to replicate *in vivo*. Both of these activities can be rescued by a high-affinity E2 binding site (pm8/BS9, Figure 3B). When tested for replication *in vitro* in the presence of competitor, the pm8 mutant is rescued poorly by the addition of E2 (3-fold; Figure 7B, compare lanes 1–3 and 4–6). When the affinity of the E2 binding site is increased (pm8/BS9), the rescue approaches the levels for the wt template (>20-fold; Figure 7B, lanes 7–12), demonstrating good correlation between *in vitro*

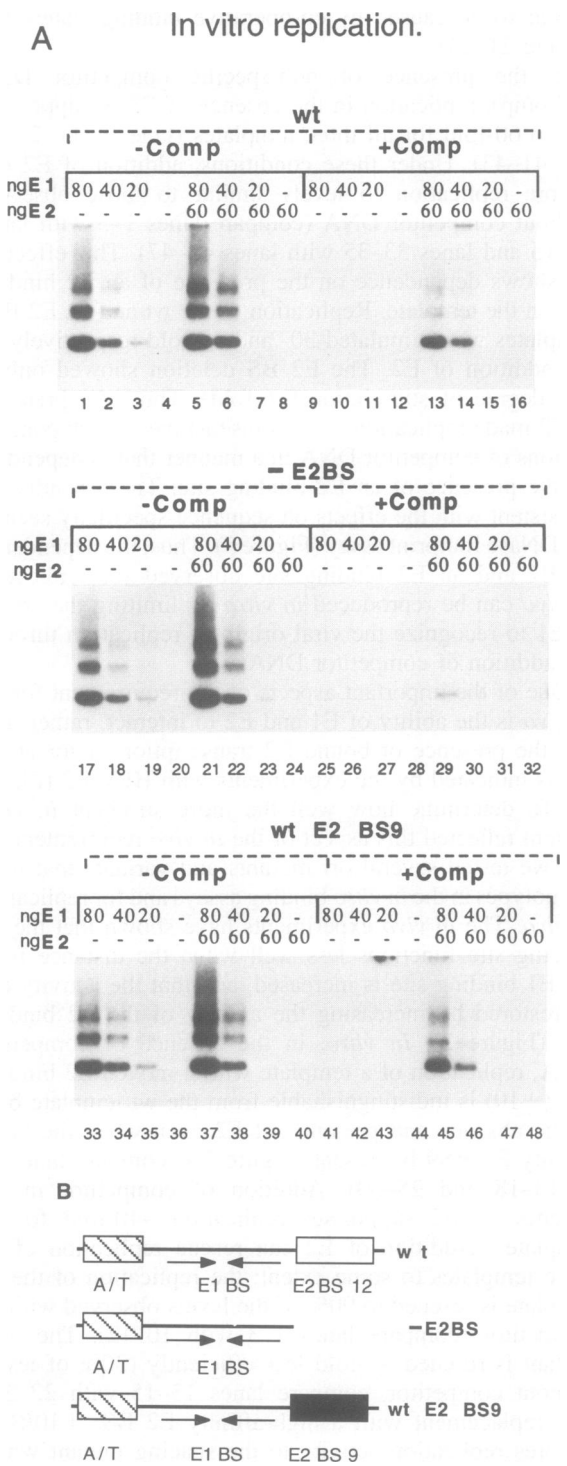

Fig. 6. Replication of BPV *in vitro* can be made responsive to E2 and an E2 binding site by the addition of competitor DNA. (**A**) *In vitro* replication reactions using three different templates were assembled as described in Materials and methods. The effect of E2 on *in vitro* replication was tested at three different E1 concentrations, either in the absence or the presence of competitor DNA as indicated in the figure. The replication products were analyzed on non-denaturing agarose gels. (**B**) Schematic drawing of the three templates used in the replication reactions in Figure 6A. All ori constructs are based on the minimal ori plasmid 7914-27 which contains the BPV sequences between nt 7914-27 cloned into the pUC polylinker between the *Xba*I and *Hin*dIII sites. The wt ori carries the wt E2 BS12, the deletion mutant (–E2 BS) lacks precisely the 12 nt that constitute the E2 binding site. In the mutant wt E2 BS9 the wt E2 BS12 was replaced with the high-affinity E2 BS9 (see Figure 2).

replication in the presence of competitor and *in vivo* replication. Finally, the HPV-11 E2 protein, which can bind to the E2 binding site but is unable to interact with BPV E1 and is inactive for replication *in vivo*, was also tested for replication in the presence of competitor DNA. As shown in Figure 7C, under conditions where BPV E2 is greatly stimulatory for replication (lanes 1–6), HPV E2 is devoid of activity (lanes 7–12).

These results demonstrate that an *in vitro* replication system can be generated that accurately reflects all aspects of the requirement for E2 that we observe for *in vivo* replication. In the presence of competitor DNA, replication *in vitro* becomes dependent on E2 and a correctly positioned E2 binding site. Importantly, the same effect can not be achieved by simply lowering the concentration of E1, demonstrating that the effect of competitor DNA is not simply to lower the effective concentration of E1. Taken together, these results indicate clearly that a critical function for E2 in replication *in vivo* as well as *in vitro*, is to contribute to the sequence specificity of the E1–E2–ori complex.

Discussion

We have demonstrated here that a key role in initiating BPV replication is played by a complex consisting of E1 and E2 bound to the origin of replication. All mutations that affect the formation of this complex, either (i) in ori affecting the E2 binding site, (ii) in ori affecting the E1 binding site or (iii) mutations that affect the spacing between the two sites, simultaneously affect replication. The most straightforward interpretation of these results is that the formation of the E1–E2–ori complex is obligatory for initiation of replication *in vivo*. Interestingly, the individual binding activities of the E1 and E2 proteins appear to be critical only in the context of the formation of the E1–E2–ori complex. Mutations in the E1 binding site that reduce the ability of E1 to bind to the ori can be compensated, both for E1–E2–ori complex formation as well as for replication, by mutations in the E2 binding site that increase the affinity for E2. E2 binding site mutants that fail to bind E2 in the absence of E1 can bind E2 in the form of the E1–E2 complex, albeit at reduced efficiency (Ustav *et al.*, 1993 and A.Stenlund, unpublished data). Likewise, increased spacing between the E1 and E2 binding sites that results in loss of both complex formation and replication activity, can be compensated for, both for complex formation and replication, by increased affinity of the E2 binding site.

The distance dependence between the E1 and E2 binding sites and the ability to rescue complex formation and replication by increasing the affinity of one site, indicates that the interaction between E1 and E2 on the ori is essential for complex formation and hence for replication. This notion is substantiated by the lack of activity of the related HPV-11 E2 protein which can bind with high affinity to the ori but fails to form the E1–E2–ori complex and also fails to support replication. The sum of these results indicates that the activities of E1 and E2 that we measure in our binding assays, i.e. the ability to interact with each other which results in co-operative binding of the two proteins, is a required activity for replication *in vivo*. As we have demonstrated, one of the results

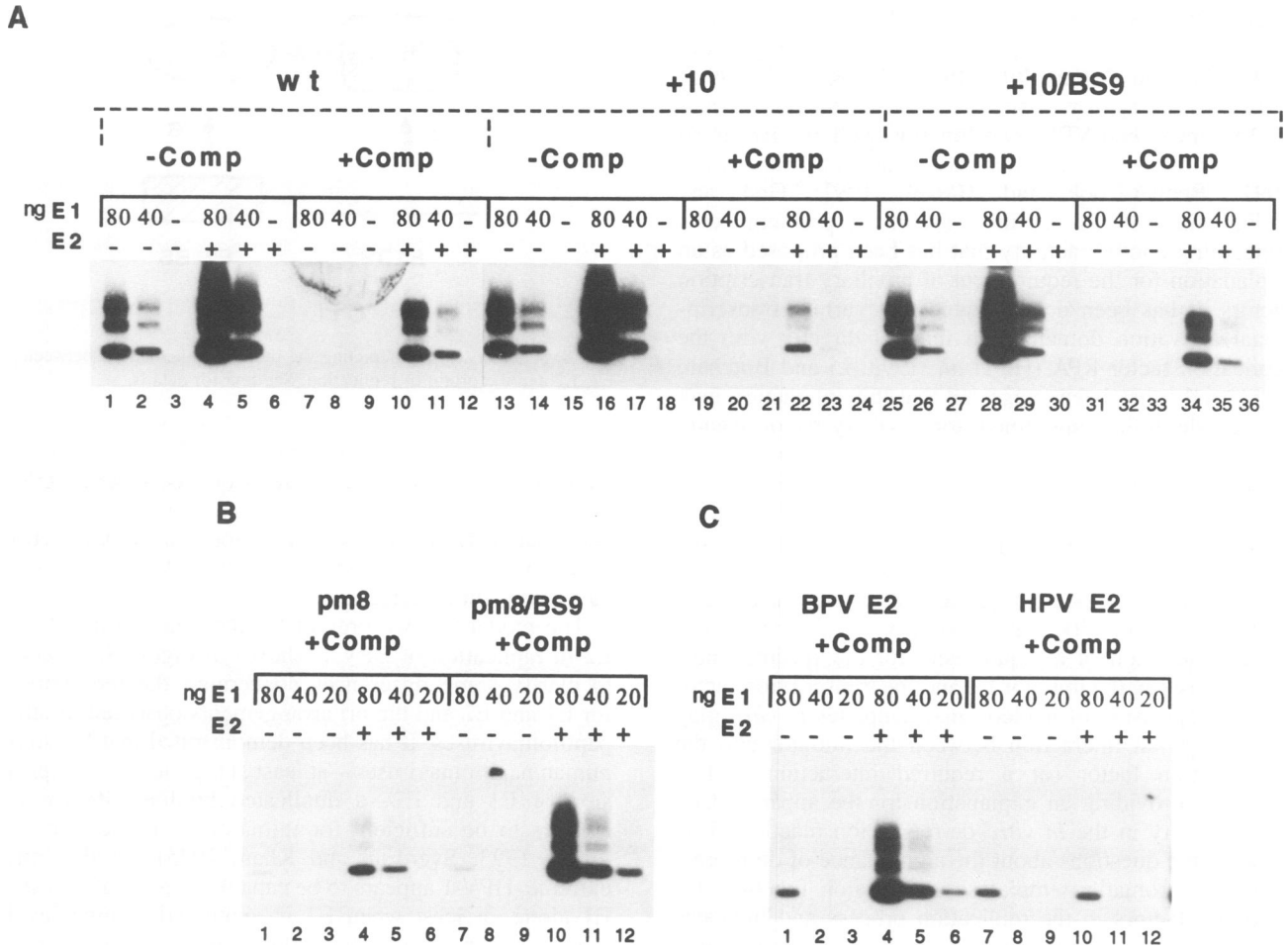

Fig. 7. E2-dependent replication *in vitro* shows similar requirements as replication *in vivo*. (**A**) The wt ori, the spacing mutant +10, and the same mutant in the context of a high-affinity E2 BS9 (+10/BS9) were tested for replication *in vitro* in the absence and presence of E2. The effect of E2 replication *in vitro* is described in Figure 6A. (**B**) A point mutation in the E1 binding site, pm8, and the same mutant in the context of the high-affinity E2 BS9 (pm8/BS9) were tested for replication *in vitro* in the presence of competitor DNA. (**C**) The wt ori template where the E2 BS had been replaced with a high-affinity HPV-11 E2 binding site was tested for the ability of BPV E2 and HPV-11 E2 to rescue *in vitro* replication in the presence of competitor as described in Figure 6.

of this interaction is a substantially increased sequence specificity when E1 and E2 bind together compared with when the two proteins bind individually. These results argue strongly that an important function of E2 is to provide the essential protein–protein interaction that is required to increase the ability of the initiator E1 to discriminate between 'specific' and 'non-specific' DNA. This conclusion is supported by the fact that by challenging the sequence specificity of E1 by introduction of non-specific competitor DNA, a standard *in vitro* replication system that shows little dependence on E2, can be made highly dependent on E2 and an appropriately positioned E2 binding site, closely reproducing the *in vivo* requirements.

Consistent with a role for E2 as a specificity factor is the fact that E1 is not a very specific DNA binding protein. While we cannot directly determine the binding constant for the formation of the E1 complex—since we do not know the exact composition of the complex—we can from the footprinting studies estimate that the K_d for E1 binding to the ori has to be at least in the nanomolar range and E1 is consequently a good DNA binder comparable with, for example, lambda repressor (Johnson *et al.*, 1979, 1980). However, the competition experiments demonstrate that the affinity for non-specific DNA is only 10- to 100-fold lower, depending on the type of competitor. In contrast, the differential between affinities for specific and non-specific DNA for lambda repressor is estimated to be 10^8-fold (Johnson *et al.*, 1981). This very small difference between the affinity for specific and non-specific DNA would appear to be inconsistent with specific ori recognition in the presence of sequences with high complexity and provides a rationale for an increased specificity requirement.

Previous *in vitro* replication studies have indicated that auxiliary factors can function in de-repression of templates assembled into chromatin (Cheng and Kelly, 1989; Cheng *et al.*, 1992; Li and Botchan, 1994). However, some experimental results are not consistent with the interpretation that this is the sole function of auxiliary factors. For example, as has recently been demonstrated in the case of BPV, nucleosome repression on reconstituted chromatin templates can clearly be overcome *in vitro* by the hybrid transactivator Gal4/VP16. However, Gal4/VP16 has no activity for replication *in vivo* and cannot replace E2 (Li and Botchan, 1994). A similar situation exists for the SV40 system: Gal4/VP16 functions for de-repression of

nucleosomal templates *in vitro* but does not function as an auxiliary factor for replication *in vivo* (Cheng *et al.*, 1992; Guo and DePamphilis, 1992; Hoang *et al.*, 1992). This does not appear to be a peculiarity specific to Gal4/VP16, since Gal4/VP16 can function well *in vivo* as an auxiliary factor for polyomavirus replication (Baru *et al.*, 1991; Bennet-Cook and Hassel, 1991; Guo and DePamphilis, 1992). A similar specificity problem exists concerning another activity that has been proposed as an explanation for the requirement of auxiliary transcription factors. It has been demonstrated that various transcriptional activation domains can interact directly with the replication factor RPA (He *et al.*, 1993; Li and Botchan, 1993) and it has been suggested that this interaction may play a role in the requirement for auxiliary factor. Again, however, Gal4/VP16—which has no activity for SV40 or BPV replication *in vivo*—interacts well with RPA.

In contrast, the stringent *in vitro* system that we have described here shows near-perfect correlation between the requirements for replication *in vivo* and *in vitro*, including the specificity for the appropriate transcription factor. The very strict specificity that we observe in this system is clearly the result of a requirement for the specific interaction between the initiator E1 and the transcription factor E2. De-repression of nucleosomal templates *in vitro* may not require an interaction between the initiator and the transcription factor (or a required interaction is less specific), providing an explanation for the apparent lack of specificity in the *in vitro* de-repression reaction. This raises some questions about the significance of de-repression of chromatin templates as a major function for auxiliary factors in the replication process and indicates that transcription factors provide another function which is more dependent on the identity of the transcription factor. The obvious explanation for the observation that certain transcription factors function for certain replicons, but not for others, is that the identity of the initiator determines which transcription factors can 'work' due to a requirement for a specific interaction between the two. It has been suggested, based on this differential specificity, that the transcription factors that function as auxiliary factors for SV40 and polyomavirus function through interaction with T-antigen (Guo and DePamphilis, 1992) just as E2 functions through interaction with E1.

One way of viewing the sum of these observations is that both the interaction with RPA as well as the de-repression of chromatin templates represent functions that are important for replication *in vivo*, but show little or no specificity, i.e. many different transcription factors can provide these functions. This leads to the conclusion that a third function of auxiliary factors is responsible for the specificity requirement that is observed *in vivo*. We have demonstrated here that one major function of the auxiliary factor E2 is to participate in the assembly of an active replication complex by increasing the binding specificity of the initiator. The auxiliary factors in other replicons could have a similar function. In adenovirus, the DNA binding domain of the transcription factor NF I can interact with and promote binding of the adenovirus preterminal protein–DNA polymerase complex to the adenovirus ori, exerting an effect on specificity (Bosher *et al.*, 1990; Chen *et al.*, 1990; Mul *et al.*, 1990; Mul and van der Vliet, 1992). It is possible that auxiliary factors perform a similar function for SV40 and polyomavirus. However, since T-antigen appears to be a more sequence-specific DNA binder than E1, ori recognition may require less of a contribution from an auxiliary factor and the interaction between the initiator and the auxiliary factor may thus have more subtle effects.

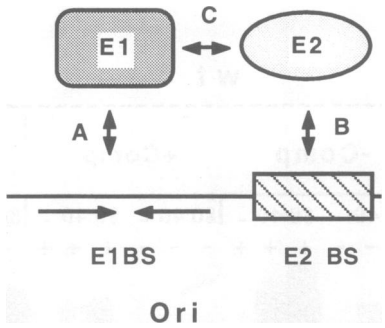

Fig. 8. Schematic figure depicting the proposed interactions between E1, E2 and the origin of replication. See text for details.

The model that we present for the function of E1 and E2 in replication of BPV is shown in Figure 8. It serves to clarify some differences concerning the requirement for E1 and E2 and the ori arrangements observed in other papillomaviruses. It has been demonstrated that for some human papillomaviruses—at least at high levels of expression of E1 and E2—a duplicated binding site for E2 appears to be sufficient for initiation of replication (Lu *et al.*, 1993; Sverdrup and Khan, 1995). At the other extreme, HPV-1 appears to be capable of replicating using E1 alone, at least when E1 is expressed at high levels (Gopalakrishnan and Khan, 1994). Both of these results can be incorporated in the model in Figure 8 by simply allowing for different contributions from the three different interactions 'A', 'B' and 'C' that we have listed. If, for example, interaction 'A' between E1 and the E1 binding site is sufficiently specific, the other two interactions may be of lesser importance, especially at high concentrations of E1(HPV-1). If, on the other hand, the interaction between E1 and DNA is very weak, a strong interaction between E2 and E2 binding sites ('B') and between E1 and E2 ('C'), would be required to allow E1 to find the ori (HPV-11). These results are consistent with a recent linker substitution analysis of the HPV-11 ori (Russell and Botchan, 1995) which indicated that the E1 binding site is of modest importance for ori activity. A very weak interaction between E1 and DNA in HPV-11 is also consistent with the recent observation that an *in vitro* replication system utilizing E1 and ori from HPV-11 requires E2 for full activity (Kuo *et al.*, 1994).

Here we have described how protein–protein interactions between DNA-binding proteins that individually display a low degree of sequence specificity, by virtue of their interaction, can generate a highly specific combined complex. This ability of the E1 and E2 proteins to co-operate and together generate highly specific DNA binding shows similarities for example with the well-studied α2 and MCM1 system from *Saccharomyces cerevisiae*. These proteins bind to interdigitated sites, and binding of MCM1 can stabilize the binding of α2 through a protein–protein interaction (Johnson, 1992; Smith and Johnson, 1992). An important difference, however, is the very relaxed

requirement for spacing and distance that exists for co-operative binding between E1 and E2. Even though activity is reduced by increased distance between the sites, this loss of activity is not absolute and can be compensated for by increasing the affinity of the E2 binding site. *In vitro*, we have difficulties detecting the E1–E2–ori complex when the spacing between the E1 and E2 binding sites is increased beyond 20 nt. However, we have previously demonstrated that multimerized E2 binding sites are functional *in vivo* from a distance of several kb (Ustav *et al.*, 1993). We assume that the function of E2 is the same whether or not E2 binding sites are located at a distance or close to the E1 binding site, and consequently that E2 can confer increased sequence specificity from a great distance.

This observation extends the similarity between the function of E2 in replication and its ability to function as a conventional transcription factor and activate transcription from a distal enhancer. Perhaps E2, and even transcription factors in general, perform essentially the same function in transcription and replication, but with different targets. In this model, instead of interacting with, and providing increased sequence specificity for the initiator E1, the transcriptional activator could interact with and provide increased sequence specificity for the TBP-containing complexes that bind to the TATA box and are responsible for the initiation of transcription.

Materials and methods

Plasmids and mutants

All BPV ori constructs and mutants were generated in a context of an ori fragment containing the BPV sequences between nt 7914–27 cloned between the *Xba*I and *Hin*dIII restriction sites in pUC 19. All point mutations in the E1 and E2 binding sites and the spacing mutations were generated by PCR using standard procedures. The E2 binding site replacement mutants were generated by PCR using a primer containing the high-affinity E2 binding site E2 BS9 from the BPV genome (Li *et al.*, 1989). The template –E2 BS used in Figure 6 contains all the sequences contained in the wt ori, lacking only the 12 nt constituting the E2 binding site.

Gel-shift assays

Gel-shift assays were carried out by mixing the probe (5000 c.p.m., 50 pg) with the specific proteins at the indicated concentrations in a buffer containing 20 mM KPO_4, pH 7.5, 0.1 M NaCl, 1 mM EDTA, 10% glycerol, 0.7 mg/ml BSA, 0.1% NP40, 2 mM DTT in a total volume of 10 µl in the presence of 20 ng of pUC 119. After incubation at room temperature for 20 min, glutaraldehyde was added to a final concentration of 0.02% and incubation was continued for another 20 min. Loading dye was added to quench the cross-linking reaction and the samples were resolved on 1% agarose gels. All probes for gel-shift assays and for DNaseI footprints were generated by PCR using end-labeled universal primers and the same templates that were used for replication assays. The probes therefore contain, in addition to the ori sequence, also polylinker sequence from pUC 19 as well as the universal primer binding sites.

DNase footprints

DNase footprint analyses were carried out using standard procedures (Jones *et al.*, 1985). The probe (10 000 c.p.m., 0.1 ng) was incubated with the indicated quantities of specific purified proteins in a buffer containing 20 mM KPO_4, pH 7.5, 0.1 M NaCl, 1 mM EDTA, 10% glycerol, 0.7 mg/ml BSA, 0.1% NP40, 2 mM DTT in a total volume of 10 µl. After incubation at room temperature for 20 min, 10 µl of a solution containing 10 mM $MgCl_2$ and 5 mM $CaCl_2$ was added and the samples were treated with DNaseI for 1 min. The digestion was terminated by the addition of STOP buffer (0.2 M NaCl, 20 mM EDTA, 1% SDS, 250 µg/ml tRNA) followed by extraction with phenol/chloroform and ethanol precipitation.

Transient replication assays

Transient replication assays were carried out essentially as described (Ustav and Stenlund, 1991). All replication assays were performed in the E1 and E2 expressing cell line 4.15, except the experiments in Figure 4B, which were performed by co-transfections of the ori plasmids with E1 and E2 expression vectors. The BPV proteins were expressed from the expression vectors pCGEag (E1) and pCGE2 (E2) (Ustav and Stenlund, 1991). The HPV-11 E1 and E2 proteins were expressed from the expression vectors pMTE1 (E1) and pMTE2 (E2) (Chiang *et al.*, 1992).

Preparation of FM3A cell extracts

The growth of mouse FM3A cells and the preparation of extracts has been essentially described by Yang *et al.* (1991). Cell cultures (8 l, 7–8×10^5 cells/ml) were harvested by centrifugation, washed with PBS and with ice-cold hypotonic buffer (20 mM HEPES–KOH, pH 7.5, 5 mM KCl, 1 mM EDTA, 0.5 mM DTT). One half of the initial packed cell volume of hypotonic buffer was added, the cells were resuspended and left to swell for 15 min on ice. The cells were disrupted in a Dounce homogenizer and the lysate was adjusted to 0.25 M NaCl and left on ice for 60 min. The lysate was centrifuged at 25 000 r.p.m. for 30 min. The supernatant was recovered and dialyzed for 6 h against 2 l of 20 mM HEPES–KOH, pH 7.5, 10 mM NaCl, 1 mM EDTA, 0.5 mM DTT with one change of buffer. The extract was centrifuged at 8000 r.p.m. for 10 min before freezing aliquots in liquid nitrogen.

In vitro replication

Standard *in vitro* replication assays were performed in 10 µl reaction mixtures containing 30 mM HEPES–KOH, pH 7.5, 7 mM $MgCl_2$, 1 mM DTT, 4 mM ATP, 0.2 mM each of GTP, UTP, CTP; 0.1 mM each of dATP, dGTP and dTTP, 10 µM of [^{32}P]dCTP (2 µCi; 800 Ci/ mmol), 40 mM creatine phosphate, 1 µg creatine kinase, 3 µl FM3A cell extract (30 µg of total protein) and template DNA. Reactions were incubated for 60 min at 37°C. The reactions were stopped by addition of SDS to 1% and EDTA to 10 mM and treated with proteinase K followed by phenol/chloroform extraction. The products were analyzed by electrophoresis on 0.8% agarose gels in TAE buffer. Template DNA for *in vitro* replication assays was purified by double banding in CsCl gradients. The concentration of template in the *in vitro* reactions was 5 ng/10 µl. Double-stranded, high-molecular weight polymer $(dA/dT\text{-}dT/dA)_n$ was used as a competitor and added to 20-fold mass excess over BPV template (100 ng/10 µl). This particular competitor was chosen for two reasons: (i) dA/dT does not serve as a template for incorporation of the labeled nucleotide (dCTP) and does not generate background labeling due to repair synthesis; (ii) dA/dT is an ~10-fold more potent competitor than dI/dC and therefore can be added to lower final concentrations. The incorporation under these conditions with the wt template was ~0.3 pmol of dCTP incorporated in 1 h using 5 ng of template and 80 ng of E1 in a volume of 10 µl. The rate of incorporation shows a linear increase with increasing concentration of template up to 10 ng/µl of reaction mixture.

Acknowledgements

We thank Winship Herr for critical reading of the manuscript and T.Sedman for excellent technical assistance. This work was supported by a grant from NIH CA 13106.

References

Baru,M., Shlissel,M. and Manor,H. (1991) *J. Virol.*, **65**, 3496–3503.
Bennet-Cook,E.R. and Hasssel,J.A. (1991) *EMBO J.*, **10**, 959–969.
Bosher,J., Robinson,E.C. and Hay,R.T. (1990) *New Biol.*, **2**, 1083–1090.
Chen,M., Mermod,N. and Horwitz,M. (1990) *J. Biol. Chem.*, **265**, 18634–18642.
Cheng,L. and Kelly,T.J. (1989) *Cell*, **59**, 541–551.
Cheng,L., Workman,J.L., Kingston,R.E. and Kelly,T.J. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 589–593.
Chiang,C.M., Ustav,M., Stenlund,A., Ho,T.F., Broker,T.R. and Chow,L.T (1992) *Proc. Natl Acad. Sci. USA*, **89**, 5799–5803.
DePamphilis,M.L. (1993) *Annu. Rev. Biochem.*, **62**, 29–63.
Gillette,T.G., Lusky,M. and Borowiec,J.A. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 8846–8850.
Gopalakrishnan,V. and Khan,S.A. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 9597–9601.
Guo,Z.-S. and DePamphilis,M.L. (1992) *Mol. Cell. Biol.*, **12**, 2514–2524.

He,Z., Brinton,B.T., Greenblatt,J., Hassel,J.A. and Ingles,C.L. (1993) *Cell*, **73**, 1223–1232.
Hoang,A.T., Weidong,W. and Gralla,J.D. (1992) *Mol. Cell. Biol.*, **12**, 3087–3093.
Holt,S.E., Schuller,G. and Wilson,V.G. (1994) *J. Virol.*, **68**, 1094–1102.
Johnson,A.D. (1992) *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 975–1006.
Johnson,A.D., Meyer,B.J. and Ptashne,M. (1979) *Proc. Natl Acad. Sci. USA*, **75**, 1783–1787.
Johnson,A.D., Pabo,C.O. and Sauer,R.T. (1980) *Methods Enzymol.*, **65**, 839–856.
Johnson,A.D., Poteete,A.R., Lauer,G., Sauer,R.T., Ackers,G.K. and Ptashne,M. (1981) *Nature*, **294**, 217–223.
Jones,K., Yamamoto,K. and Tjian,R. (1985) *Cell*, **42**, 559–572.
Kuo,S.-R., Liu,J.-S., Broker,T.R. and Chow,L.T. (1994) *J. Biol. Chem.*, **269**, 24058–24065.
Li,R. and Botchan,M.R.(1993) *Cell*, **73**, 1207–1221.
Li,R. and Botchan,M.R. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 7051–7055.
Li,R., Knight,J., Bream,G., Stenlund,A. and Botchan,M. (1989) *Genes Dev.*, **3**, 510–526.
Lu,J.Z.-J., Sun,Y.-N., Rose,R.C., Bonnez,W. and McCance,D.J. (1993) *J. Virol.*, **67**, 7131–7139.
Lusky,M. and Fontane,E. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 6363–6367.
Lusky,M., Hurwitz,J. and Seo,Y.-S. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 8895–8899.
Le Moal,M.A., Yaniv,M. and Thierry,F. (1994) *J. Virol.*, **68**, 1085–1093.
McBride,A.A., Romanczuk,H. and Howley,P. (1991) *J. Biol. Chem.*, **266**, 18411–18414.
Mohr,I.J., Clark,R., Sun,S., Androphy,E.J., MacPherson,P. and Botchan,M. (1990) *Science*, **250**, 1694–1699.
Mul,Y.M. and van der Vliet,P.C. (1992) *EMBO J.*, **11**, 751–760.
Mul,Y.M., Verrijzer,C.P. and van der Vliet,P.C. (1990) *J. Virol.*, **64**, 5510–5518.
Muller,F., Seo,Y.-S. and Hurwitz,J. (1994) *J. Biol. Chem.*, **269**, 17086–17094.
Park,P., Copeland,W., Yan,L., Wang,T., Botchan,M. and Mohr,I.J. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 8700–8704.
Russell,J.A. and Botchan,M.R. (1995) *J. Virol.*, **69**, 651–660.
Seo,Y.-S., Muller,F., Lusky,M. and Hurwitz,J. (1993a) *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) *Proc. Natl Acad. Sci. USA*, **90**, 2865–2869.
Smith,D.L. and Johnson,A.D. (1992) *Cell*, **68**, 133–142.
Spalholz,B.A., McBride,A.A., Sarafi,T. and Quintero,J. (1993) *Virology*, **193**, 201–212.
Sverdrup,F. and Khan,S.A. (1995) *J. Virol.*, **69**, 1319–1323.
Ustav,M. and Stenlund,A. (1991) *EMBO J.*, **10**, 449–457.
Ustav,M., Ustav,E., Szymanski,P. and Stenlund,A. (1991) *EMBO J.*, **10**, 4321–4329.
Ustav,E., Ustav,M., Szymanski,P. and Stenlund,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 898–902.
Wilson,V.G. and Ludes-Meyers,J. (1991) *J. Virol.*, **65**, 5314–5322.
Yang,L., Li,R., Mohr,I., Clark,R. and Botchan,M.R. (1991) *Nature*, **353**, 628–633.
Yang,L., Mohr,I., Fouts,E., Lim,D.A., Nohaile,M. and Botchan,M. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 5086–5090.

Received on July 3, 1995; revised on September 11, 1995