The EMBO Journal vol.10 no.2 pp.449-457, 1991

Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames

Mart Ustav[1] and Arne Stenlund

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[1]On leave of absence from the Laboratory of Oncogenesis, Estonian Biocenter, Tartu University, Tartu, Estonia, USSR

Communicated by U.Pettersson

Bovine papillomavirus (BPV) DNA is maintained as an episome with a constant copy number in transformed cells and is stably inherited. To study BPV replication we have developed a transient replication assay based on a highly efficient electroporation procedure. Using this assay we have determined that in the context of the viral genome two of the viral open reading frames, E1 and E2, are required for replication. Furthermore we show that when produced from expression vectors in the absence of other viral gene products, the full length E2 transactivator polypeptide and a 72 kd polypeptide encoded by the E1 open reading frame in its entirety, are both necessary and sufficient for replication BPV in C127 cells.
Key words: DNA replication/papillomavirus/transactivator

Introduction

The bulk of knowledge about DNA replication stems from studies performed on prokaryotes including prokaryotic viruses and episomal elements where both the basic enzymatic machinery and regulatory devices have been studied extensively. A regulatory feature that has attracted particular interest is how the copy number of the genetic material is controlled and maintained. In several cases, in particular for certain prokaryotic plasmids, a very detailed understanding of copy number control and maintenance now exists (Nordstrom *et al.*, 1984; Nordstrom and Austin, 1989). These systems have in common the feature of a single, or very few, origins of replication per genetic element. In eukaryotes, the much larger genomes necessitate multiple sites for initiation of replication, which complicates both the regulation of replication and the analysis thereof. A considerable amount of information about the basic enzymatic machinery and the process of initiation of replication has been learned from the study of lytic mammalian viruses (Stillman, 1989; Challberg and Kelly, 1989). These viruses, which are exceptions to the multiple origin of replication rule of higher eukaryotes, in most cases lack the kind of clearly defined replication control, that the host cell requires, and hence are unsuitable for study of the host temporal and copy number control systems. The exceptions are two completely unrelated viruses, Epstein-Barr virus (EBV) and Bovine papillomavirus (BPV). Both of these viruses have a latent stage where the viral DNA replicates only during S phase as a multicopy nuclear plasmid with a regulated copy number in the cell, and appear to be stably inherited (Mecsas and Sugden, 1987).

Replication of BPV has been studied mainly in two different types of assays. In stable transformation assays a stable cell line is selected and the state of the viral DNA (i.e. if the viral DNA is episomal or integrated) is determined. This assay is quite stringent and measures not only DNA replication but also stability and segregation. A transient assay is more direct since replicated material is analyzed early after transfection in the whole population of cells. DNA is considered to have replicated if it has acquired resistance to the restriction enzyme *Dpn*I which will not cut DNA replicated in mammalian cells because of the lack of a specific methylation (Peden *et al.*, 1980). Stable replication assays have yielded conflicting results concerning the gene requirements for BPV replication. The E2 open reading frame has, in some studies, been shown not to be required for replication (Lusky and Botchan, 1985; Lusky and Botchan 1986a), while others have found an absolute requirement for E2 (Sarver *et al.*, 1984; Groff and Lancaster, 1986; Rabson *et al.*, 1986; DiMaio and Settleman, 1988). All published accounts agree that mutations in the E1 open reading frame affect replication. One study demonstrates that N-terminal and C-terminal mutants in E1 can complement each other for replication, indicative of two complementation groups in E1 (Lusky and Botchan 1986b), while on the other hand, some studies have failed to show this complementation (Lambert and Howley, 1988; Schiller *et al.*, 1989). The requirements for replication have also been described using a transient replication assay (Lusky and Botchan, 1986a; Lusky and Botchan, 1986b, Berg *et al.*, 1986). Interestingly, in these studies the requirements for transient replication differed from the requirements for stable replication. Certain mutations in the N-terminal part of the E1 open reading frame that all failed to replicate in the stable assay were reported to replicate in the transient assay. This apparent discrepancy was explained by a model where the establishment of BPV as a stably replicating plasmid was postulated to take place in two stages and that a defect in any of the functions required for either of the two stages would be scored as replication deficient in the stable assay, while only defects in the first stage would affect the transient assay. This two stage model also suggested that two distinct types of replication were required to establish BPV as a stably replicating multicopy plasmid. The initial stage being an amplification of the incoming DNA to the final copy number, replication of viral DNA would have to proceed at a greater rate than replication of the cellular DNA. The second stage would involve maintenance of the copy number and consequently replication at the same rate as the cellular DNA.

In an attempt to settle these unresolved issues, we optimized an electroporation procedure (Neumann *et al.*, 1982; Potter *et al.*, 1984; Chu *et al.*, 1987) to get DNA uptake in a large fraction of the transfected cells. We then employed this procedure for transient replication assays. In this paper we use this assay to define for the first time the

two viral factors that together are necessary and sufficient for viral DNA replication. One of these factors is the viral transactivator E2, which appears to be required in a capacity other than the hitherto reported role as a regulator of BPV transcription (Spalholz *et al.*, 1985; Spalholz *et al.* 1987; Lambert *et al.*, 1987; Stenlund and Botchan, 1990). The other factor is a 72 kd polypeptide that is encoded by the E1 open reading frame in its entirety.

Results

Replicating BPV DNA can be detected transiently in C127 cells provided that the transfection efficiency is sufficiently high

Based on the average copy number of BPV DNA in stably replicating cells (~ 100), we estimated that with a detection level of $1-10$ pg we would require in the order of 10^3-10^4 cells replicating the BPV DNA to detect replication. To measure our transfection efficiency we transfected a β-galactosidase expressing plasmid, pON 260 (Spaete and Mocarski, 1985) and stained the cells *in situ* with X-gal 36 h after transfection. Using the originally described polybrene transfection method (Kawai and Nishizawa, 1984; Lusky and Botchan, 1986b), our transfection efficiency was quite low ($0.01-0.1\%$ of the cells) in spite of our attempts to optimise the system. Therefore, this method would require very large numbers of cells as starting material to get a detectable signal (10^6-10^8 cells). In contrast, by using electroporation under appropriately optimized conditions we found that $10-20\%$ of the surviving cells stained positive for β-galactosidase. We chose therefore the electroporation method to investigate BPV replication.

For the experiment shown in Figure 1, 5×10^6 C127 cells were transfected by electroporation with increasing quantities of wildtype BPV DNA. The cells from each electroporation were plated to ten 100 mm dishes and low molecular weight DNA from one plate was harvested every 24 h using an alkaline lysis procedure. The DNA was digested with *Dpn*I which cuts only methylated (unreplicated) DNA (Peden *et al.*, 1980) and *Hin*dIII which will linearize BPV DNA. The products were analyzed on agarose gels followed by Southern transfer and hybridization to a BPV specific probe. As shown in Figure 1 a prominent *Dpn*I resistant band with the size of linear BPV appeared between 24 and 48 h after transfection and increased over time. The amount of *Dpn* sensitive material appears proportional to the quantity of input DNA, while the amount of replicated material in the upper band does not increase substantially at levels of input DNA above 1 μg. We interpret this to mean that the quantity of DNA taken up per cell increases with increasing amounts of input DNA, but that the number of cells taking up and replicating the input DNA does not substantially change. Below input levels of 0.5 μg the quantity of replicated material is strictly proportional to the amount of input DNA (data not shown). This is consistent with the idea that at low DNA concentrations only very few molecules of DNA are delivered to each cell and that DNA replication is responsible for eventually raising the copy number. What is noticeable for the lower concentrations of input DNA is that the rate of accumulation of DNA is biphasic. The quantity of replicated DNA after the first 48 h appears to increase only at a rate that is accounted for by the increase in cell number, and that, per cell, the copy number of BPV DNA remains constant, whereas the rate of increase is rapid earlier in the experiment.

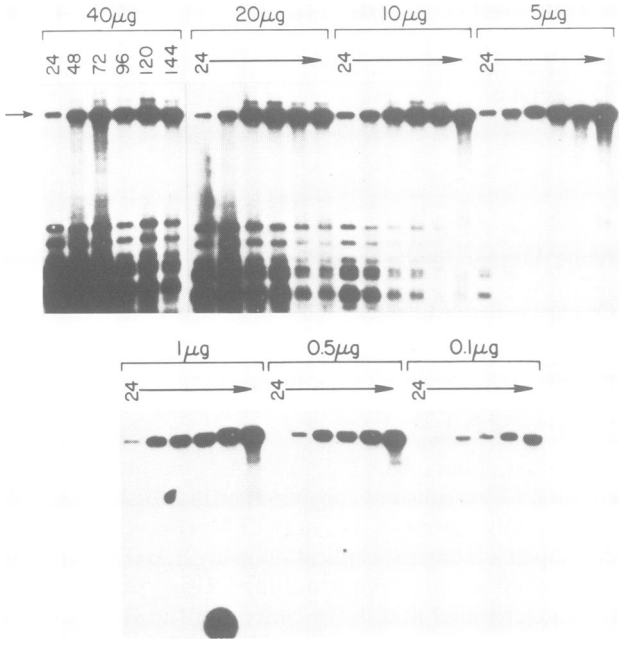

Fig. 1. Transient replication of BPV DNA in C127 cells. BPV DNA was excised from the plasmid pmlBPV by cleavage with *Bam*HI and the linear DNA was used for electroporation of C127 cells. Cells were harvested every 24 h and low molecular weight DNA was digested with *Hind*III and *Dpn*I. The *Dpn* resistant material (replicated DNA) is linearized by *Hind*III and forms an 8 kb band at the top of the gel (arrow). The *Dpn* sensitive material (unreplicated DNA) forms a series of lower molecular weight bands.

To assess further the fidelity of the replication assay we performed transformation assays using the identical electroporation conditions as for the replication assays. In several independent experiments we obtained on the average $4-5 \times 10^4$ foci/μg of BPV DNA. This number is ~ 2 orders of magnitude higher than that generally obtained by polybrene transfections. The transformation follows single hit kinetics and in that respect is identical to the results that have been obtained with viral infections (Dvoretsky *et al.*, 1980). Thus, there is good correlation between easily detectable transient replication and very high focus forming frequency.

Gene products from the E1 and E2 open reading frames are required for replication

To determine which gene products were required for transient replication we assayed a series of mutations throughout the BPV early region (Figure 2a). The frameshift mutations in E6, E7 and E5 replicated in a manner indistinguishable from the wildtype genome, while a frameshift in the E1 open reading frame, and a translation termination linker (TTL) in the E2 open reading frame that also affects the reading frames E3 and E4 failed to replicate to a detectable level. To verify that these mutations affected functions required in *trans*, we performed complementation experiments. The E2 TTL mutant which did not replicate by itself was cotransfected with a wildtype genome which was separated from the plasmid sequences to give a different size linear molecule. As shown in Figure 2b, in the presence of the wildtype genome the E2 TTL mutant replicated as did the

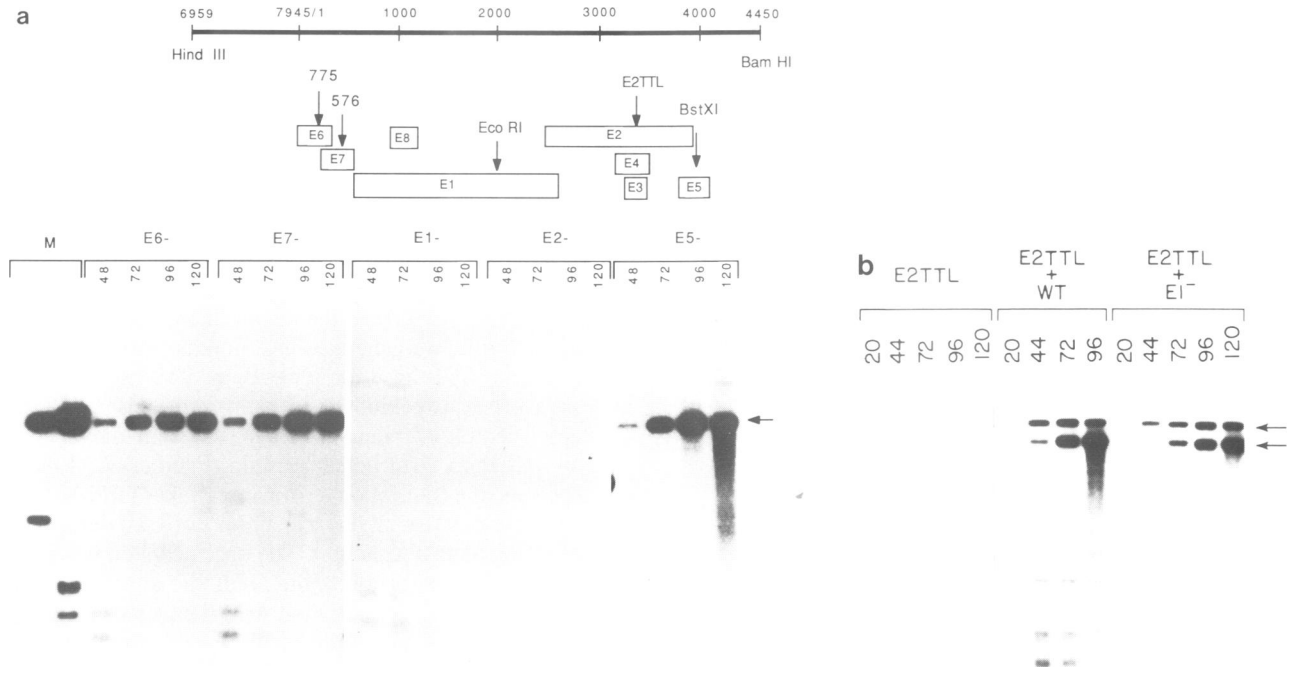

Fig. 2. Mutants in two of the early open reading frames are deficient for transient replication. (**a**) A schematic figure of the early region of BPV between the unique *Hind*III and *Bam*HI sites with the eight early open reading frames shown as boxes. The arrows denote the location and the names of the mutations that were tested in the assay below. 1 μg of the viral genome with the mutations indicated in the upper panel were assayed for replication as described in Figure 1. The mutants in E6, E7 and E5 replicate at wildtype levels. The E1⁻ and E2⁻ mutants show no detectable replication. M is a marker with 100 and 200 pg of linear BPV DNA. (**b**) Complementation of the replication deficient mutants restores replication. The E2⁻ mutant E2 TTL which was replication deficient was co transfected with either the wildtype genome or the E1⁻ mutant. To distinguish between the two genomes in each assay, the wildtype genome and the E1 mutant were transfected separated from the prokaryotic vector pML, while E2TTL was still linked to the vector. This makes the size difference between the DNAs approximately 2.7 kb.

wildtype genome, indicating that the E2TTL mutation affected a function required in *trans*. The E1 mutant could be similarly complemented by the E2 TTL mutant, indicating first that the E1 mutation affected a *trans* function, and also that the E1 and E2 mutations fall into different complementation groups. An additional point that we have noted repeatedly is that the upper band, which represents the viral DNA that is linked to the prokaryotic plasmid DNA, appears to replicate less well than the viral genome which was separated from the plasmid sequences. This inhibition was also apparent in focus assays, where the efficiency of focus formation was considerably greater when the plasmid sequences were removed.

Replication competence can be restored to an E2 mutant through complementation with an E2 expression vector

The E2 TTL mutation affects several known and potential gene products. This part of the E2 open reading frame includes sequence present in the three known forms of E2 polypeptides, and in addition, in this region the E3 and E4 open reading frames overlap with E2 (Choe *et al.*, 1989; Lambert *et al.*, 1989; Lambert *et al.*, 1990). To determine which of these gene products were required for transient replication we performed a series of complementation assays using expression vectors that we constructed to produce the three defined forms of E2. The structure of these expression vectors is shown schematically in Figure 3a. pCGE2 encodes a 48−50 kd protein from the entire E2 reading frame, pCGE2C encodes a 30 kd protein that is initiated from an internal ATG in E2 and that is identical to the C-terminal part of E2 and pCGE8/E2 expresses a 28 kd form of E2 that splices together 11 amino acids from the E8 open reading frame with the C-terminal 204 amino acids of the E2 open reading frame. To ensure that these constructs expressed the appropriate polypeptides at comparable levels, we transfected the different expression vectors into 293 cells and analyzed the polypeptides expressed transiently by metabolic labeling followed by immunoprecipitation with a polyclonal E2 antiserum. As shown below the diagrams in Figure 3a the constructs produced polypeptides of the expected size. When tested for complementation of the E2TTL mutant in the replication assay only pCGE2, which encodes the full length form of E2, could complement the defect in replication (Figure 3b). pCGE2C and pCGE8/E2 lacked detectable activity. Furthermore, the activity of the pCGE2 expression vector was dependent on the C-terminal part of the E2 open reading frame, since a deletion that removes the C-terminal 33 amino acids abolished complementation (data not shown). These results indicate clearly that the E2 polypeptide is required for replication. We cannot formally rule out that the E3 or E4 open reading frames also play a role in replication, we think however that it is unlikely that the expression vector produces these polypeptides. From the functional analyses that have been performed on E2 proteins, it is known that the different polypeptides appear to have different functions (Lambert *et al.*, 1987; Hubbert *et al.*, 1988; Choe *et al.*, 1989). The full length form of the protein is a transcriptional activator whereas the two shorter forms appear to be competitive repressors of this activity. Some of the functional regions of the E2 polypeptides have been mapped (Giri and Yaniv, 1988; Dostatni *et*

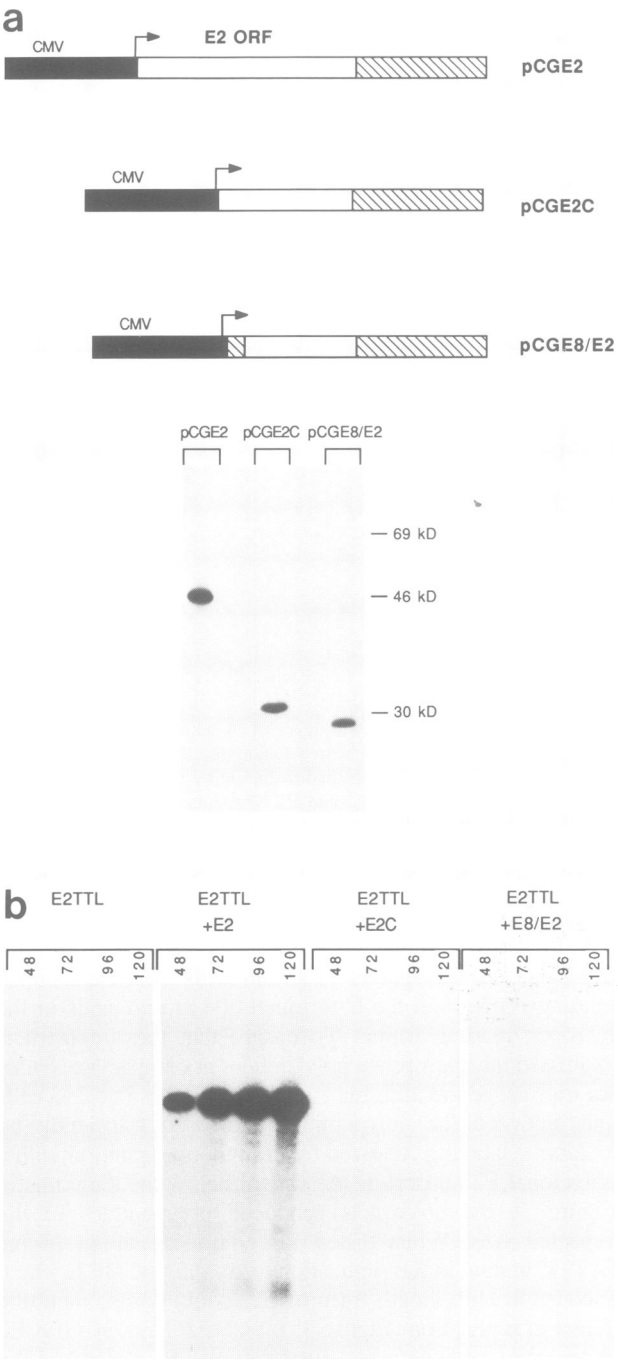

Fig. 3. Complementation of the E2 TTL mutant with E2 expression vectors. (**a**) Schematic drawings of the E2 expression vectors used in this experiment. pCGE2 contains the whole E2 open reading frame. pCGE2C contains a C-terminal fragment from E2 encoding 250 amino acids. pCGE8/E2 encodes 11 amino acids from the E8 open reading frame spliced to the C-terminal 204 amino acids of the E2 open reading frame. Below is shown the products for each of the expression vectors after transient expression in 293 cells and immunoprecipitation with a polyclonal E2 antiserum. The molecular weights for the polypeptides are ~48–50 kd (pCGE2), 30–31 kd (pCGE2C), 27–28 kd (pCGE8/E2). (**b**) 1 μg of the E2 TTL mutant was co-transfected with 5 μg of each of the three E2 expression vectors pCGE2, pCGE2C and pCGE8/E2 shown in Figure 3b and assayed for replication as described in Figure 1.

al. 1988; McBride *et al.*, 1988; McBride *et al.*, 1989). Starting from the C-terminal part of the protein, dimerization and DNA binding require only 100 C-terminal amino acids. These properties are shared by all forms of E2 proteins. The N-terminal portion of the open reading frame is required for the transactivation function, and consequently, the two shorter forms of E2 lack this function. This indicated that the dimerization and DNA binding functions that are present in the repressor forms were not sufficient to complement the replication defect, but that some function specific to the transactivator was required.

The E1 open reading frame is required in its entirety for replication

To characterize further the requirement for the E1 open reading frame, we generated insertion mutations throughout E1 and assayed these for replication. The insertion mutations were of two kinds, either, in-frame insertions of an oligonucleotide encoding 13 amino acids of a known epitope from influenza virus hemagglutinin (Field *et al.*, 1988), or insertions of a translation termination linker (TTL). The in-frame insertions, with one exception, were all defective for replication (data not shown). The one exception was the 'N-terminal-most' insertion at the *Sma*I site at amino acid 33 (Sma-Epi), which replicated at wild type levels (Figure 4). To determine if this region of the E1 open reading frame encodes the E1 replication function we inserted a translation termination linker at the same *Sma*I site (SmaTTL). This mutant was defective for replication, indicating that translation through this region of the reading frame is required. This led us to the conclusion that the replication function was encoded by the E1 open reading frame in its entirety. This notion was further corroborated through complementation assays between the SmaTTL mutant and the other nonfunctional E1 insertion mutants, none of which gave rise to replication (data not shown). These results confirmed that the N-terminal and the C-terminal parts of the E1 open reading frame belong to the same complementation group, and probably encode one polypeptide. Since this result was in clear disagreement with results presented previously, we obtained one of the protoptye mutants from the earlier studies (Lusky and Botchan 1986a; Berg *et al.*, 1986). This mutant, E1-Sma, is a frameshift mutation at the *Sma*I site, and was reported to replicate in a transient assay (Berg *et al.*, 1986; Lusky and Botchan, 1986a). As shown in Figure 4, this mutant does not replicate in our assay, consistent with the results we obtained with the SmaTTL mutant. We also inserted a translation termination linker at an *Nru*I restriction site 9 nucleotides upstream of the first ATG codon in E1. As shown in Figure 4, this mutation (NruTTL) did not affect replication, indicating that the E1 polypeptide is either initiated at the N-terminal ATG codon, since this is the only in-frame ATG upstream of the *Sma*I site, or that coding sequence from upstream could be spliced in between the *Nru*I and *Sma*I sites.

One of our objectives was to construct an E1 expression vector that could complement E1 mutants. Since nothing is known about which promoter is responsible for synthesis of an E1 mRNA or what the structure of such an mRNA is, we used as much sequence as possible upstream of the reading frame to ensure that any putative upstream exon would be included in the construct. To generate the expression vector we inserted the coding sequence for E1 plus 1.4 kb of sequence upstream of the E1 open reading frame into the expression vector pCG (Tanaka and Herr, 1990). This expression vector (pCGMlu, Figure 5a) was capable of complementing the SmaTTL mutant (data not shown).

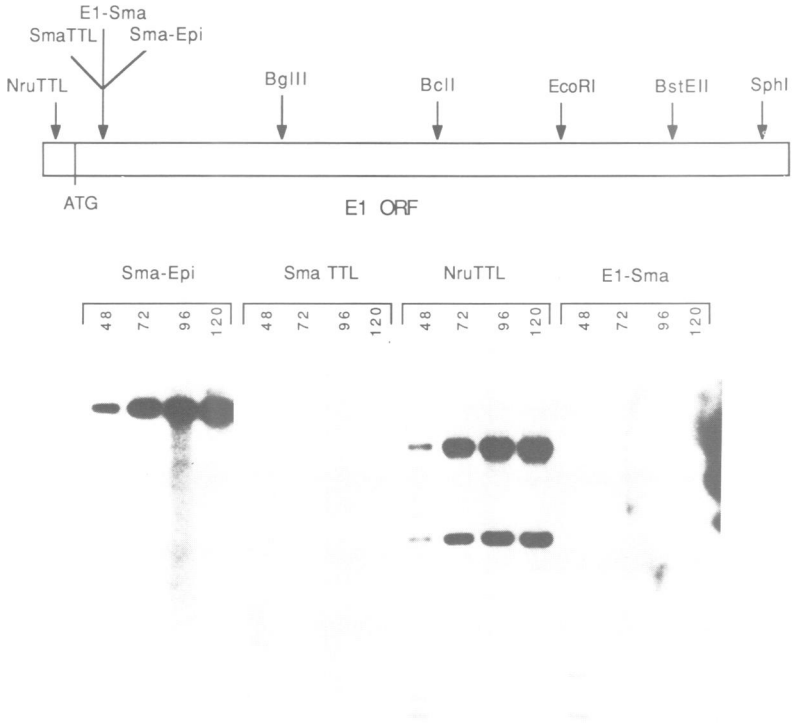

Fig. 4. A schematic figure of the E1 open reading frame with the position of the various mutations that were tested shown by arrows. The NruTTL mutation is located immediately upstream of the first ATG of the E1 open reading frame. Three different mutations at the *Sma*I site were generated; a termination linker insertion (SmaTTL), and in frame insertion of 13 amino acids from influenza hemagglutinin (Sma-Epi), and a frame shift generated by fill in of the *Xma*I site (E1-Sma). The remaining mutants are all insertions of the hemagglutinin epitope. Below is shown replication assays using four of the mutants shown above. 1 μg of each mutant was transfected and analyzed as described in Figure 1, except that the harvested samples were digested with *Xba*I instead of *Hin*dIII. The presence of two bands in the *Nru*TTL lanes is due to the presence of an *Xba*I site in the TTL linker, in addition to the single site in wildtype BPV.

We then constructed several derivatives of this vector, pCGMlu/Sma-Epi which contains the epitope insertion at the *Sma*I site, pCGMlu/SmaTTL which contains the frame shift mutation at the *Sma*I site and pCGMl/NruTTL which contains the translation termination linker at the *Nru*I site immediately upstream of the first ATG codon in E1. We also modified the expression vector by removing most of the sequence upstream of the open reading frame, placing the promoter at nucleotide 619 (pCGEag). In this construct the sequence between the promoter and the first ATG in E1 lacks potential initiation codons in any reading frame. This expression vector was functional and could complement an E1 mutant (Figure 5a). Consequently an upstream exon required for E1 function could be ruled out.

To verify the genetic data we transfected the E1 expression vectors into COS cells and performed immunoprecipitations after metabolic labeling. The immunoprecipitations were carried out using either a polyclonal serum raised against the C-terminal portion of the E1 open reading frame, or a monoclonal antibody specific to the hemagglutinin epitope, pCGMlu, pCGMlu/NruTTL and pCGEag all produced a polypeptide ~72 kd that could be precipitated with the polyclonal antiserum but not with the monoclonal antibody (Figure 5b). pCGMlu/Sma-Epi produced a slightly larger polypeptide (74 kd) that could be precipitated with both the polyclonal serum and the monoclonal antibody against the hemagglutinin epitope. pCGMlu/SmaTTL produced no specific polypeptide detectable with either antiserum. These results show clearly that the 72 kd polypeptide that we detect is encoded by the E1 open reading frame, since the monoclonal antibody against the hemagglutinin epitope will detect the polypeptide only when the epitope has been inserted into the coding sequence of E1. The slight change in molecular weight of the protein due to the epitope insertion is consistent with this conclusion. Furthermore, the 72 kd polypeptide appears to be encoded in its entirety from the E1 open reading frame without additional coding sequence requirements, since the pCGEag vector that lacks upstream initiation codons can produce this polypeptide. Based on the good correlation between the presence of the 72 kd polypeptide and replication activity, we conclude that the replication function on E1 is contributed by the 72 kd full length translation product from the E1 open reading frame. This agrees well with the recent detection of an ~68 kd polypeptide encoded from the E1 open reading frame that has been identified in BPV transformed cells (Sun *et al.*, 1990).

E1 and E2 proteins expressed from expression vectors are sufficient to support replication of a BPV ori fragment

Having established the requirements for viral gene products in the context of the viral genome, we asked if these gene products were also sufficient for replication. We also asked if the requirement for E2 could be restricted to the context of the viral genome, due to a requirement of E2 for E1 expression. We co-transfected the E1 and E2 expression vectors together with 3.2 kb fragment from the BPV genome from nucleotide 4780 to 80, which contains the upstream regulatory region from the early region of BPV but is devoid

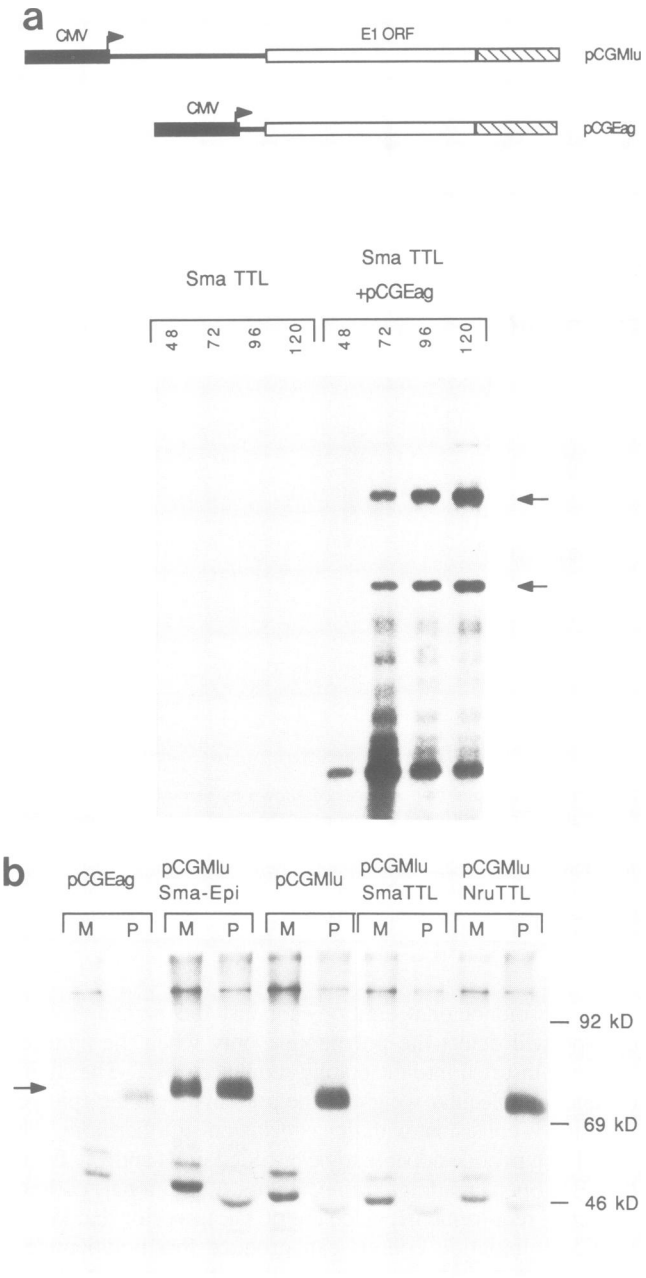

Fig. 5 (**a**) A schematic drawing showing the structure of the two prototype E1 expression vectors that were constructed based on the results from the mutagenesis of the E1 open reading frame. Both constructs contain sequences from *Bst*XI to *Bam*HI (nt3881−nt4450, stippled box) linked to either an *Mlu*I to *Avr*II fragment (nt7352−nt2766, solid line and open box) or an *Eag*I to *Avr*II fragment (nt619−nt2766, solid line and open box) inserted into the pCG vector (see Materials and methods). Below is shown a replication assay using 1 μg of the SmaTTL mutant alone or cotransfected with 5 μg of the pCGEag expression vector. Analysis was performed as described in Figure 1, except that the samples were digested with *Xba*I instead of *Hin*dIII. The additional band is due to the presence of an *Xba*I site in the TTL linker inserted at the *Sma*I site. (**b**) Immunoprecipitations from transient expression experiments in COS cells using the E1 expression constructs and their derivatives. Immunoprecipitations from each construct were carried out with two different antisera, one rabbit polyclonal serum raised against an E1-trpE fusion (P) and a monoclonal antibody specific for the hemagglutinin peptide sequence present in the Sma-Epi insertion (M). pCGMlu/Sma-Epi contains the in-frame insertion of the sequence for the influenza hemagglutinin epitope at the *Sma*I site, pCGMlu/SmaTTL contains a translation termination linker insertion at the *Sma*I site, and pCGMlu/NruTTL contains a translation termination linker insertion at the *Nru*I site in E1.

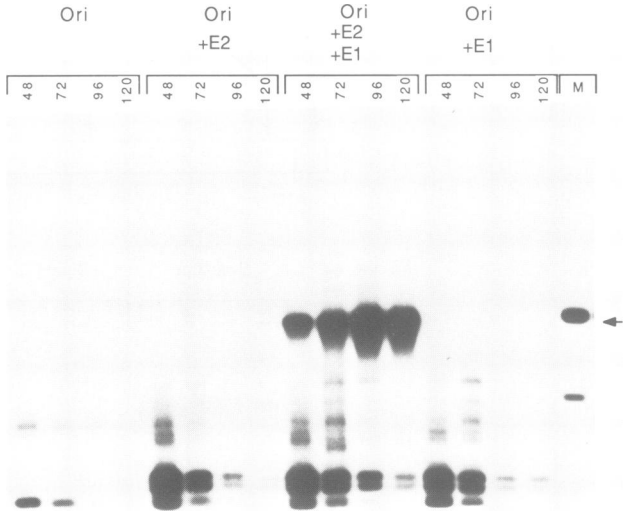

Fig. 6. A transient replication assay using the E1 and E2 expression vectors as the only source of viral *trans* factors. 2 μg of a 3.2 kb ori fragment (BPV nt 4780−80) was transfected either alone or together with 5 μg of either the pCGE2 or pCGEag expression vectors, or with 5 μg each of both expression vectors.

of early coding sequence. As shown in Figure 6, transfection of this fragment alone or with either of the expression vectors does not give rise to detectable replication, while transfection of all three DNAs together results in replication. Finally, to ensure that the gene requirements were the same in the absence of the viral genome, we tested the E2C and the E8/E2 expression vectors together with the E1 expression vector for replication. Consistent with the earlier complementation results, these proteins are inactive for replication and cannot replace the transactivator form of E2 (data not shown). Also, since a polypeptide other than the full length form of E1 has been reported to be encoded from the N-terminal part of E1 (Thorner *et al.*, 1988), and the short open reading frame E8 overlaps with the E1 open reading frame, we generated point mutations in the pCGEag expression vector to ask if these overlapping functions had any significance for replication. In pCGEag1197 the potential initiation codon ATG at nucleotide 1204 of the E8 open reading frame is changed to ACG and in pCGEag1235 the splice donor sequence at nucleotide 1235 that has been reported to be used for the short N-terminal polypeptide from E1 (Thorner *et al.*, 1988) has been mutated. Both of these mutants supported replication in a manner indistinguishable from the wildtype expression vector pCGEag (data not shown). This shows clearly that the 72 kd polypeptide from the E1 open reading frame together with the 48 kd polypeptide from the E2 open reading frame are the only viral *trans* factors required for replication.

Discussion

From the results presented here it appears that previous difficulties in obtaining transient replication of BPV in C127 cells have been due mainly to low transfection efficiencies. Transient replication appears quite efficient, based on the assumption that the copy number does not exceed what is seen in transformed cells (~100). A very large fraction of the cells that take up viral DNA will also replicate it. The 50−100 pg of replicated DNA that we detect in a timepoint

in a typical transient replication assay (for example see Figure 2a), corresponds to $5 \times 10^4 - 1 \times 10^5$ cells with an average copy number of 100 molecules/cell. The total number of cells in the timepoint is $\sim 1 \times 10^6$, which means that 5–10% of the cells replicate BPV transiently. This number is very similar to our measured frequency of DNA uptake, which means that virtually every cell that takes up BPV DNA will also replicate it. Furthermore, the cells that take up and replicate BPV DNA also appear to progress and become stably transformed at a very high frequency indicated by the high focus forming frequency.

Previous work on the requirements for transient replication of BPV shows several differences with the results presented in this paper (Lusky and Botchan, 1986; Berg *et al.*, 1986). No requirement for the E2 open reading frame was found in the earlier studies, whereas we see an absolute dependence on a specific form of E2 for replication. Also, the results concerning the E1 open reading frame differ quite significantly. In the earlier work the E1 open reading frame was defined as consisting of two complementation groups one of which (R) was required for transient and stable replication, the other (M) being dispensable for transient replication, but required for stable replication. We find no evidence for a second replication function in E1 and our results indicate very clearly that the replication function required for transient replication (as well as for stable replication) is encoded by the entire E1 reading frame. The prototype E1-Sma mutant that replicated transiently in the earlier study did not replicate in our transient assay (Figure 4). The origin of these differences in results is hard both to explain and to explore further. As mentioned previously we cannot reproduce the transfection procedure that was used in the earlier studies, so a direct repetition of those experiments under identical conditions is not possible. We have pursued the more obvious possibilities by testing C127 cells from several different sources, which all behave identically in this regard. The idea from the earlier studies that the transiently replicating material represented a specific early stage of BPV replication, as was initially suggested, appears to have very little basis since no direct link was ever found between the stable and transient replication assays. On the contrary, for the transient assays using the polybrene procedure, the focus forming frequency for example was reported to be no higher than a few hundred foci/μg, and it is unknown what the fate of the transiently replicating DNA was, since it clearly did not end up in transformed cells (Lusky and Botchan, 1986b).

The very good correlation between the requirements for replication in transient assays and in stable assays, and also the good correlation between efficient transient replication and focus formation has changed our perception of what occurs during the early stage of BPV replication. We do not believe that our transient assay reflects a specific stage of BPV establishment. Rather it seems that the entire process of establishment is included within the assay with amplification taking place rapidly at the earliest time points. In general, at 72 h post transfection only the increase in replicated DNA that is expected from normal cell doubling can be detected, which by definition indicates maintenance. It is noteworthy that we see no sign of a differential gene requirement for the two stages of replication. The mutations that we have tested fall into only two classes, competent and defective for replication. The replication competent mutants all behave identically which indicates that the amplification and maintenance stages of replication, to the limits of resolution of our assay, require the same gene products. This would suggest that a conventional immediate-early to late-early switch in viral gene expression is not responsible for the transition from amplification to maintenance. This idea is also supported by analysis of viral gene expression under replication conditions which show no change in the activity of any of the viral promoters throughout the amplification and maintenance stages (P.Szymanski and A.Stenlund, unpublished observation).

These observations lead us to believe that what has been termed the amplification and maintenance stages are not that different. The change from replication faster than the cellular DNA to replication at the same rate as the cellular DNA could be achieved simply by a copy number control function, where the level of an essential factor which is kept constant, determines the final copy number. In this model amplification would occur as long as the factor is in excess relative to the copy number, while maintenance would result when the factor becomes limiting. One candidate for the limiting factor is the E2 polypeptide since some experiments suggest that the level of E2 is related to the copy number of BPV DNA both in transient assays (M.Ustav and A.Stenlund, unpublished observation) and in transformed cells (Lambert *et al.*, 1990; Riese *et al.*, 1990).

We have identified for the first time the two viral polypeptides that are required for BPV replication. The E1 polypeptide is poorly characterized, but an activity of the E1 polypeptide in replication has been predicted by the weak but significant homology of the E1 open reading frame to large T-antigen from SV 40 (Clertant and Seif, 1984). Presumably, the role of E1 will be similar to that of T-antigen, which carries out various biochemical functions that are required directly in the replication process. The E2 polypeptide is relatively well characterized as a transactivator, nevertheless the role of E2 in the replication process is more difficult to predict. In the context of the viral genome the requirement of a transactivator is not unexpected, for example, expression of E1 from its promoter could very well be E2 dependent. Surprisingly the requirement for E2 persists even when E1 is supplied from a source that is not E2 responsive, and no other viral genes are present, pointing to a more direct role of E2 in replication. One explanation for the E2 requirement would be that the E2 polypeptide, which is known to activate an E2 dependent enhancer in the upstream regulatory region (URR), is required for replication because the BPV origin of replication is dependent on this enhancer for function. This would be analogous to the enhancer requirement for polyoma virus DNA replication (DePamphlis, 1988), and would also be very similar to the role that EBNA-1 appears to play in the replication of EBV from *ori*P. Replication from oriP is dependent on EBNA-1, and multiple binding sites for EBNA-1 that are located on the *ori*P fragment (Rawlins *et al.*, 1985; Lupton and Levine, 1985; Mecsas and Sugden, 1987). These binding sites have been shown to act as an EBNA-1 dependent enhancer for transcription (Reisman *et al.*, 1985; Reisman and Sugden, 1986). Interestingly, the fragment that we have used as an *ori* fragment from BPV contains the URR region which carries multiple E2 binding sites and also is an E2 dependent enhancer (Androphy *et al.*, 1987; Haugen *et al.*, 1987; Hawley-Nelson *et al.*, 1988; Li *et al.*, 1989; Spalholz *et al.*, 1985; 1987). This role would also be consistent with the fact that only the transactivator form of E2 can support replica-

tion and that the DNA binding domain of E2 is required for the replication function. This however, may not be the only function of E2, since under certain circumstances replication is independent of the E2 enhancer but still requires E2 (M.Ustav and A.Stenlund, unpublished observation). The E2 polypeptide might thus play additional roles in the replication process, apart from its function as a transactivator.

Materials and methods

Plasmid constructs

The BPV genome was used in the form of pML BPV (Lusky and Botchan, 1985) which consists of the viral genome linearized at the unique *Bam*HI site and inserted into the *Bam*HI site of the plasmid pML.

Mutants in the viral genome. The TTL mutations at *Nru*I (nt 838), *Sma*I (nt 945) and E2 TTL (nt 3351) were generated by insertion of a 14 nucleotide *Xba*I linker (New England Biolabs) that contains termination codons in all three reading frames. The mutation *Bst*XI in E5 is a frameshift mutation generated by digestion with *Bst*XI, removal of overhangs with T4 DNA polymerase and religation. The mutations in E1 at the *Sma*I, *Bgl*II, *Bcl*I, *Eco*RI *Bst*EII and *Sph*I sites were generated by insertion of an oligonucleotide encoding 13–15 amino acids from influenza virus hemagglutinin. The mutants in E6 (775) and E7 (576) and one of the mutations in E1 (E1-Sma) have been described previously (Lusky and Botchan, 1985; Berg *et al.*, 1986).

E2 expression vectors. The E2 expression vectors were generated by insertion of subgenomic fragments from the E2 open reading frame into the expression vector pCG (Tanaka and Herr, 1990). The fragments were inserted between the restriction sites *Xba*I downstream of the TK leader sequence and *Bgl*II in the second exon of the β-globin sequence, by addition of an *Xba*I linker to the upstream end of each fragment and fusing the *Bam*HI site at nucleotide 4450 in the BPV sequence to the *Bgl*II site in the vector, pCGE2 contains the BPV sequence between nucleotide 2600 and 4450. In this construct a point mutation was generated at nucleotide 3092 (ATG–ATC) which changes the initiation codon for E2C to an Ile codon. pCGE2C contains the sequence between 3089 and 4450. pCGE8/E2 contains the insert from a cDNA clone with the sequence from nucleotide 1010–1235 joined by a splice to the sequence from 3225–4450. In all these constructs the first ATG is the authentic initiation codon.

E1 expression vectors. The E1 expression vector pCGMlu was constructed by first deleting the sequence between *Avr*II (nt 2766) and *Bst*XI (nt 3881) from the BPV genome. An *Xba*I linker was inserted at *Mlu*I (nt 7352) and the *Xba*I (nt 7352) to *Bam*HI (nt 4450) fragment was inserted between the *Xba*I and *Bgl*II sites in the vector pCG. pCGEag was constructed from pCGMlu by deletion of the sequence between *Mlu*I (nt 7352) and *Eag*I (nt 619) upstream of the E1 open reading frame. The mutant pCGEag 1197 carries a point mutation at nt 1205 that changes an ATG codon (Met) to an ACG codon (Thr) in the E8 open reading frame. The mutant pCG Eag 1235 carries a point mutation at nucleotide 1235 that changes the splice donor sequence AGGTA to AAGTA.

Cell culture

The cell lines C127, COS, NIH 3T3 and 293 were all maintained in DME + 10% fetal bovine serum.

Determination of transfection efficiency was carried out by *in situ* staining of transfected cells 24–36 h after transfection with the β-galactosidase expressing plasmid pON 260. Plates were washed thoroughly with PBS, the cells were fixed with 0.05% glutaraldehyde for 15 min, washed again with several changes of PBS and stained with a staining solution containing 0.05% X-gal, 0.2 mg/ml spermidine 10 mg/ml potassium ferrocyanide, 8 mg/ml potassium ferricyanide and 3mM $MgCl_2$ in PBS. The plates were examined in the microscope after overnight incubation at 37°C.

Replication assays

C127 cells were trypsinized, centrifuged and resuspended in DME + 10% fetal bovine serum (FBS) + 5 mM BES pH 7.2 at a density of 2×10^7 cells/ml and stored at room temperature, 0.25 ml of cell suspension was mixed with plasmid DNA, 50 μg of salmon sperm DNA was added as carrier and the mixture was pipetted into a disposable electroporation cuvette. The cell/DNA mixture was subjected to an electric discharge (210 V, 960 μF) using a BioRad Gene Pulser with a capacitance extender. After the discharge the cell/DNA mixture was left at room temperature for 10–15 min and then layered on top of 10 ml of DME + 10% FBS in a centrifuge tube. The cells were pelleted at low speed, resuspended in 10 ml of DME + 10% FBS and half the cells were plated onto five 100 mm dishes. A time point was taken every 24 h and the cells were lysed by using a modified version of the alkaline lysis procedure as described in Maniatis *et al.*, 1982. Briefly, the plate was washed with PBS, 600 μl of a 1:2 mixture of solution I and solution II was added and the plate was incubated on ice for 5 min. 300 μl of solution III was added and after 10 min on ice the lysate was scraped off into a microfuge tube. The supernatant was recovered after centrifugation in the cold, and was precipitated with 0.6 vol of isopropanol. After centrifugation the samples were dissolved in 200 μl of a solution containing 20 mM Tris–HCl pH 8.0, 100 mM NaCl, 10 mM EDTA, 0.2% SDS and 200 μg/ml proteinase K and incubated at 50°C for 1 h. The samples were extracted with phenol/chloroform and precipitated with ethanol. After a second precipitation with ethanol the sample was dissolved in 20 μl of TE buffer containing 20 μg/ml RNase A and incubated at 68°C for 20 min. 10 μl of restriction mixture containing 6 units of *Dpn*I and 6 units of *Hind*III was added and incubated overnight at 37°C. The samples were run on 0.8% agarose gels in TAE buffer and blotted to nitrocellulose using standard conditions. High specific activity BPV probes were generated by random priming of pMLBPV DNA using a kit from Amersham and hybridizations were carried out at 65°C in 6×SSC, 5× Denhardt's solution, 0.5% SDS, 200 μg/ml salmon sperm DNA, and 10 mM EDTA.

Quantitation of replication assays: To determine the absolute amount of hybridizing DNA, known quantities of plasmid DNA were included as markers on the gels. After hybridization and exposure to X-ray film, bands were cut out from the nitrocellulose filter and counted in a liquid scintillation counter.

Focus assays were carried out using the conditions described above with the exception that 3–50 ng of BPV DNA was used. The cells were plated on ten 100 mm dishes and foci were counted three weeks after transfection.

Immunoprecipitations

For transient detection of the E2 polypeptides 5×10^6 293 cells were transfected by electroporation with 5 μg of the respective expression construct. Half of the cells were plated, and 24 h after the transfection the cells were labeled with 100 μCi of a methionine–cysteine mixture in methionine/cysteine free medium. After 3 h a lysate was prepared by boiling the cells in RIPA buffer (0.5% NP 40, 0.5% Tween 20, 0.5% deoxycholic acid, 0.15 M NaCl, 10 mM KCl, 20 mM Tris–HCl pH 7.5 and 1 mM EDTA) with 1% SDS. After a 10-fold dilution in RIPA to reduce the SDS concentration to 0.1% the lysate was precleared by addition of pre-immune serum and protein A–Sepharose. Precipitations were carried out with a polyclonal rabbit antiserum raised against a trpE-E2 produced in *E.coli* (T.Talpsepp and M.Ustav, unpublished) and the precipitated products were analyzed on 12% SDS–polyacrylamide gels (Laemmli, 1970).

For transient detection of E1, COS cells were transfected wth 5 μg of expression construct by electroporation. The cells were plated and 36 h after transfection they were labeled with 500 μCi of Tran^{35}S–label for 4 h. Cells were washed with PBS and lysed on the plate by addition of 1 ml of RIPA. Insoluble material was pelleted by centrifugation at 10 000 *g*, the supernatant was pre-cleared by addition of non-specific IgG and protein A– Sepharose and precipitations were carried out with a polyclonal rabbit antiserum raised against a C-terminal E1/trpE fusion (T.Talpsepp and M.Ustav, unpublished) and with a monoclonal antibody 12CA5 raised against influenza hemagglutinin peptide HA1(75-110) (Field *et al.* 1988).

Acknowledgements

We thank Bruce Stillman and Paul Szymanski for comments on the manuscripts and Masafumi Tanaka also for helpful discussions throughout this work. We also thank Tiit Talpsepp for preparation of antisera against the E1 and E2 polypeptides and Urmas Saarma for generation of the point mutations in E1. This work was supported by a grant from the National Institutes of Health CA 13106-19.

References

Androphy,W.J., Lowy,D.R. and Schiller,J.S. (1987) *Nature*, **325**, 70–73.
Berg,L., Lusky,M., Stenlund,A. and Botchan,M. (1986) *Cell*, **46**, 753–762.
Challberg,M.D. and Kelly,T.J. (1989) *Annu. Rev. Biochem.*, **58**, 671–718.
Choe,J., Vaillancourt,P., Stenlund,A. and Botchan,M. (1989) *J. Virol.*, **63**, 1743–1755.
Chu,G., Hayakawa,H. and Berg,P. (1987) *Nucleic Acids Res.*, **15**, 1311–1326.

Clertant,P. and Seif,I. (1984) *Nature*, **311**, 276–279.
DePamphlis,M.L. (1988) *Cell*, **52**, 635–638.
DiMaio,D. and Settleman,J. (1988) *EMBO J.*, **7**, 1197–1204.
Dostatni,N., Thierry,F. and Yaniv,M. (1988). *EMBO J.*, **7**, 3807–3816.
Dvoretsky,I., Shober,R., Chattopadhyay,S.K. and Lowy,D.R. (1980) *Virology*, **10**, 369–375.
Field,J., Nikawa,J.-I., Broeke,D., Macdonald,B., Rodgers,L., Wilson,I.A., Lerner,R.A. and Wigler,M. (1988) *Mol. Cell. Biol.*, **8**, 2159–2165.
Giri,I. and Yaniv,M. (1988) *EMBO J.*, **7**, 2823–2829.
Groff,D.E. and Lancaster,W.D. (1986) *Virology*, **150**, 221–230.
Haugen,T.H., Cripe,T.P., Ginder,G.D., Karin,M. and Turek,L.P. (1987) *EMBO J.*, **6**, 145–152.
Hawley-Nelson,P., Androphy,E.J., Lowy,D.R. and Schiller,E.J. (1988). *EMBO J.*, **7**, 525–531.
Hubbert,N.L., Schiller,J.T., Lowy,D.R. and Androphy,E.J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 5864–5868.
Kawai,S. and Nishizawa,M. (1984) *Mol. Cell Biol.*, **4**, 1172–1174.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Lambert,P.F., Spalholz,B.A. and Howley,P.M. (1987) *Cell*, **50**, 69–78.
Lambert,P.F. and Howley,P.M. (1988) *J. Virol.*, **62**, 4009–4015.
Lambert,P.F., Hubbert,N.L., Howley,P.M. and Schiller,J.T. (1989) *J. Virol.*, **63**, 3151–3154.
Lambert,P.F., Monk,B.C. and Howley,P.M. (1990) *J. Virol.*, **64**, 950–956.
Li,R., Knight,J., Bream,G., Stenlund,A. and Botchan,M. (1989) *Genes Dev.*, **3**, 510–526.
Lupton,S. and Levine,A.J. (1985) *Mol. Cell Biol.*, **5**, 2533–2542.
Lusky,M. and Botchan,M. (1985) *J. Virol.*, **53**, 955–965.
Lusky,M. and Botchan,M. (1986a) *Proc. Natl Acad. Sci. USA*, **83**, 3609–3613.
Lusky,M. and Botchan,M. (1986b) *J. Virol.*, **60**, 729–742.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning—A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McBride,A.A., Byrne,J.C. and Howley,P.M. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 510–514.
McBride,A.A., Schlegel,R. and Howley,P.M. (1988) *EMBO J.*, **7**, 533–539.
Mecsas,J. and Sugden,B. (1987) *Annu. Rev. Cell. Biol.*, **3**, 87–108.
Neumann,E., Schaefer-Ridder,M., Wang,Y. and Hofschneider,P.H. (1982) *EMBO J.*, **1**, 841–845.
Nordstrom,K., Molin,S. and Light,J. (1984) *Plasmid*, **12**, 71–90.
Nordstrom,K. and Austin,S.J. (1989) *Annu. Rev. Genet.*, **23**, 37–69.
Peden,K.W.C., Pipas,J.M., Pearson-White,S. and Nathans,D. (1980) *Science*, **209**, 1391–1396.
Potter,H., Weir,L. and Leder,P. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 7161–7165.
Rabson,M.S., Yee,C., Yang,Y.C. and Howley,P.M. (1986) *J. Virol.*, **60**, 624–634.
Rawlins,D., Milman,G., Hayward,S.D. and Hayward,G.S. (1985) *Cell*, **42**, 859–868.
Reisman,D. and Sugden,B. (1986) *Mol. Cell Biol.*, **6**, 3838–3846.
Reisman,D., Yates,J.L. and Sugden,B. (1985) *Mol. Cell Biol.*, **5**, 1822–1832.
Riese,D.J., Settleman,J., Neary,K. and DiMaio,D. (1990) *J. Virol.*, **64**, 944–949.
Sarver,N., Rabson,M.S., Yang,Y.C., Byrne,J.C. and Howley,P.M. (1984) *J. Virol.*, **52**, 377–388.
Schiller,J.T., Kleiner,E., Androphy,E.J., Lowy,D.R. and Pfister,H. (1989) *J. Virol.*, **63**, 1775–1782.
Spaete,R.R. and Mocarski,E.S. (1985) *J. Virol.*, **56**, 135–143.
Stenlund,A. and Botchan,M. (1990) *Genes Dev.*, **4**, 123–136.
Stillman,B. (1989) *Annu. Rev. Cell Biol.*, **5**, 197–245.
Spalholz,B.A., Lambert,P.F., Yee,C.L. and Howley,P.M. (1987) *J. Virol.*, **61**, 2128–2137.
Spalholz,B.A., Yang,Y.-C. and Howley,P.M. (1985) *Cell*, **42**, 183–191.
Sun,S., Thorner,L., Lentz,M., MacPherson,P. and Botchan,M. (1990) *J. Virol.*, **64**, 5093–5105.
Tanaka,M. and Herr,W. (1990) *Cell*, **60**, 375–386.
Thorner,L., Bucay,N., Choe,J. and Botchan,M. (1988) *J. Virol.*, **62**, 2474–2482.

Received October 15, 1990; revised on November 26, 1990