Identification of a 68-Kilodalton Nuclear ATP-Binding Phosphoprotein Encoded by Bovine Papillomavirus Type 1

SHAW SUN, LAUREN THORNER, MICHAEL LENTZ, PAUL MACPHERSON, AND MICHAEL BOTCHAN*

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

Received 2 May 1990/Accepted 18 July 1990

E1 is the largest open reading frame (ORF) of bovine papillomavirus type 1 (BPV-1) and is highly conserved among all papillomaviruses, maintaining its size, amino acid composition, and location in the viral genome with respect to other early genes. Multiple viral replication functions have been mapped to the E1 ORF of BPV-1, and evidence suggested that more than one protein was encoded by this ORF. We previously identified a small protein (M) whose gene consists of two exons, one encoded by the 5' end of the E1 ORF. We show here that a 68-kilodalton (kDa) phosphoprotein made from the E1 ORF can be detected in BPV-1-transformed cells, and we present evidence that this protein is encoded by sequences colinear with the entire E1 ORF. The full-length E1 protein immunoprecipitated from virally transformed cells and identified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis comigrates with a protein expressed from a recombinant DNA construct capable of producing only the complete E1 protein. In addition, two different antisera directed against polypeptides encoded from either the 3' or the 5' end of the E1 ORF both recognize the full-length E1 product. A mutation converting the first methionine codon in the ORF to an isoleucine codon abolishes BPV-1 plasmid replication and E1 protein production. Consistent with the notion that this methionine codon is the start site for E1, a mutant with a termination codon placed after the splice donor at nucleotide 1235 in E1 produces a truncated protein with the molecular mass predicted from the primary sequence as well as the previously identified M protein. When visualized by immunostaining, the E1 protein expressed in COS cells is localized to the cell nucleus. A high degree of similarity exists between the BPV-1 E1 protein and polyomavirus and simian virus 40 large-T antigens in regions of the T antigens that bind ATP. We show by ATP affinity labeling that the E1 protein produced in COS cells binds ATP and that this activity is abolished by a point mutation which converts the codon for proline 434 to serine. Furthermore, this mutation renders the viral genome defective for DNA replication, suggesting that the ATP-binding activity of E1 is necessary for its putative role in viral DNA replication. We conclude that the E1 ORF encodes at least two bona fide viral polypeptides in BPV-1-transformed cells: a small protein with an apparent molecular mass of 23 kDa translated from the 5' portion of E1 and a large protein with an apparent molecular mass of 68 kDa.

Bovine papillomavirus type 1 (BPV-1) provides an attractive model system for analyzing the mechanisms regulating eucaryotic DNA replication, since it can transform mouse fibroblast cells and maintain itself stably as a multicopy nuclear plasmid (14, 31, 33). This latent replication, whereby a constant complement of BPV-1 plasmids is maintained in dividing cells, is believed to be analogous to the normal viral replication in the fibroblast components of bovine warts (26). A biphasic model for BPV-1 replication in mouse C127 cells has been proposed to account for the finding that a single BPV-1 virus or plasmid can infect these cells and then amplify its copy number to 100 to 200 plasmids after only a few cell doublings (3, 14, 31). This model suggests that initially the viral DNA replicates faster than the DNA of the cell, allowing for accumulation of plasmids (establishment stage), before a switch is made to a regulated mode of replication (maintenance stage) which allows for maintenance of a constant copy number.

Little is known about the proteins actually responsible for these processes. Mutations throughout the E1 open reading frame (ORF) have been shown to eliminate one or both stages of BPV-1 replication and suggest that more than one *trans*-acting factor is encoded by this ORF (3, 34, 35). In order to understand the roles of these factors in BPV-1

replication, we have been interested in first identifying and then characterizing the protein products of the E1 ORF. Using polyclonal antisera that recognize a protein segment from the 5' portion of the E1 ORF, we previously identified a 23-kilodalton (kDa) phosphoprotein in BPV-1-transformed cells that is believed to be the product of the modulator gene (*M*) (57). This protein is thought to be encoded by *M* because its presence or absence in various E1 mutants is consistent with its having the structure proposed for an *M* protein as defined genetically and because it has the same apparent molecular mass as a product translated *in vitro* from the predicted *M* cDNA (8).

Earlier genetic experiments from our laboratory showed that mutations in the 3' and 5' portions of the E1 ORF define two complementing functions necessary for BPV-1 replication (3, 35). In the absence of any biochemical studies, it was impossible to conclude whether these two functions were the result of two distinct proteins or whether they were merely separate domains of one large protein. For this reason, in our attempt to identify an E1 protein containing sequences from the 3' portion of the E1 ORF we considered the existence of a protein encoded by the entire ORF. The conservation of the size and composition of the intact E1 ORF among all the many cloned papillomaviruses sequenced to date strongly suggested the existence of a full-length E1-encoded product (2, 5). In addition, among all of the BPV-1 cDNAs that have been identified, the only ones which could encode a protein

* Corresponding author.

with information from the 3' portion of E1 would necessarily encode a full-length E1 protein (8, 26). Moreover, in contrast to our studies with regard to the phenotypes and complementation between 5' and 3' E1 mutants, two independent studies found no evidence for separate genes. Lambert and Howley (27) as well as Schiller et al. (47) have found that BPV-1 plasmids containing mutations throughout the E1 ORF all exhibit higher transforming activity compared with the wild-type plasmid. Cotransfection of these mutants with wild-type BPV-1 or a recombinant plasmid designed to express the intact E1 ORF reverses the high-transformation phenotype of the mutants, showing that they can be complemented *in trans*. However, not only do mutants in the 3' and 5' portion of E1 exhibit the same phenotypes in this assay, but pairwise crosses of the mutants do not complement each other to give wild-type numbers of foci. This suggests that in these assays the mutants affect the same function and disrupt the gene for the same protein. Why we find reduced transformation frequencies with apparently identical mutations and can complement 5' and 3' E1 ORF mutants is not resolved. All these results together indicate that the ORF is genetically complex and that multiple proteins may be affected.

In order to define the conditions necessary to identify a putative E1 protein, we first overproduced this protein from the adenovirus major late promoter in COS cells and used extracts from these cells to refine the immunoprecipitation procedure that allowed us to identify the M protein (57). We report that a protein encoded by the entire E1 ORF does indeed exist in BPV-1-transformed C127 cells, that it is phosphorylated, and that it has an apparent molecular mass of 68 kDa, as predicted from the primary sequence of the E1 ORF. Examination of cell lines containing BPV-1 plasmids with translation-termination mutations located at different sites throughout the E1 ORF shows that they produce either no protein or a truncated E1 protein of the size predicted by the position of the mutation. The COS cell expression system also enabled us to begin characterization of the E1 protein. We show that the E1 protein is localized to the cell nucleus by using immunostaining experiments. This finding is consistent with the observations that the amino-terminal end of E1 contains a stretch of amino acids resembling nuclear localization signals found in the simian virus 40 (SV40) and polyomavirus large-T antigens (11, 18, 19, 28, 29, 42, 53). We also show that the E1 protein binds ATP in affinity-labeling experiments using reactive [α - 32 P]ATP and that a substitution of serine for proline at amino acid position 434 prevents this binding. These features, nuclear localization and ATP binding, are consistent with the possibility that the E1 protein plays a role in BPV-1 replication.

MATERIALS AND METHODS

Construction of recombinant plasmids. Expression plasmids were constructed with the vector pMT21 (see Fig. 1A), which was kindly provided by R. Kaufman (Genetics Institute). This plasmid is essentially identical to pMT2 (46), except that the string of G's preceding the dihydrofolate reductase (*dhfr*) gene has been deleted. pMT-E1 (see Fig. 1A) contains a fragment of BPV-1 sequence from nucleotides (nt) 838 to 3004 cloned into pMT21 and was constructed as follows. pMLBPV₁₀₀ (33) was cut with *Nru*I and *Acc*I, the overhangs were filled in with the Klenow fragment of DNA polymerase I, and the gel-purified 2.2-kilobase fragment was ligated into pMT21 at the *Eco*RI site, which had also been made blunt with the Klenow fragment. pMT-Repy contains a

fragment of BPV-1 sequence from nt 1515 to 2766 fused on the 3' end to a smaller fragment of BPV-1 from nt 4396 to 4425 and on the 5' end to a fragment of murine leukemia virus sequence encoding the start methionine and two amino acids of the *gag* gene product. pMT-Repy was constructed by first inserting a *Bgl*III-*Bam*HI fragment (nt 1515 to 4450) from a pMLBPV₁₀₀ plasmid that already had deletions in sequences between nt 2766 and 4396 into the *Bam*HI site of plasmid DOL⁻ (23) (kindly provided by I. Lemishka). The *Bam*HI site in this vector is located at the beginning of the MLV *gag* gene present in DOL⁻ (not shown) so that fragments inserted in frame are efficiently translated by using the normal translation sequences and initiation codon for *gag*. The BPV-1 DNA fused to the beginning of the murine leukemia virus *gag* gene was then excised from this plasmid on an *Ava*II fragment (nt 533 of the murine leukemia virus sequence to nt 4425 of the BPV-1 sequence) which was made blunt with the Klenow fragment and inserted into the *Eco*RI site of pMT21, which had also been made blunt with the Klenow fragment. pMT-M contains a fragment of the BPV-1 cDNA N7-1 sequence (8) from nt 838 to 3351 cloned into the *Bam*HI site of pMT21. The N7-1 cDNA extends from nt 138 to 4203 and contains two splices, one from nt 304 to 528 and the other from nt 1235 to 3225. The 523-base-pair *Nru*I-*Stu*I fragment from N7-1 believed to encode M had been previously inserted into the *Bam*HI site of pAR3040 (43) to allow expression by T7 RNA polymerase (not shown). This fragment was removed from pAR3040 by digestion with *Bam*HI, *Eco*RI linkers were attached, and it was inserted into the *Eco*RI site of pMT21.

Mutations were made in the E1 ORF by the gapped duplex DNA method essentially as described by Kramer and Fritz (25). The 1168-base-pair *Sma*I-*Eco*RI fragment (nt 945 to 2113) and the 1342-base-pair *Eco*RI-*Kpn*I fragment (nt 2113 to 3456) of BPV-1 cloned separately into M13mp18 (59) served as templates for mutant oligonucleotides. The 1M-I mutation at nt 851 (G→A) changes the first Met codon of E1 (ATG) to Ile (ATA; changed nucleotide is underlined); this mutation also changes a Gly (GGC) in E7 to a Ser (AGC) three amino acids before the end of the ORF. The 118oc mutation at nt 1200 (G→T) changes a Glu (GAA) in E1 to an ochre stop codon (TAA); this mutation is 4 nt upstream of the first ATG in E8 and replaces one nonconsensus nucleotide (G) with another (T) according to the optimal sequence for initiation by eucaryotic ribosomes (24), which favors C at this position. In the E8 reading frame, the 118oc mutation (CTG→CTT) leaves this position a Leu. The 128am mutation at nt 1230 to 1232 changes a Leu (CTC) in E1 to an amber stop codon (TAG) at the last codon before the 1235 splice donor in E1; this mutation changes two codons in E8, but there is no change in the protein sequence. The 131am mutation at nt 1239 (C→T) changes a Gln (CAG) in E1 to an amber stop codon (TAG) at the first codon downstream of the 1235 splice site; this mutation changes a TAC codon to a TAT codon in E8, but both encode Tyr. The 434P-S mutation at nt 2148 (C→T) changes a Pro (CCT) in E1 at amino acid position 434 to a Ser (TCT). Mutant fragments were used to replace the corresponding wild-type fragment in pMLBPV₁₀₀. In addition, the 434P-S mutation was cloned into pMT-E1 by substituting a *Bgl*III-*Bst*EII fragment (nt 1515 to 2405) carrying this mutation for the wild-type fragment in pMT-E1. The E1Sma and i2113-2 mutant DNAs have been described previously (3, 34). pNEO5' (33) contains the neomycin resistance gene of Tn5 expressed from a retroviral long terminal repeat promoter inserted into pML (32).

Cell culture and transfections. Mouse C127 cells (14) and derived cell lines carrying mutant BPV-1 plasmids, ID13 cells (31), V1216 cells (57), and monkey COS cells (15) were grown in Dulbecco modified Eagle medium (GIBCO Laboratories)—10% fetal bovine serum (J. R. Scientific) supplemented with penicillin and streptomycin.

Mutant BPV-1 plasmids were transfected into C127 cells with either Polybrene (21) or calcium phosphate (58). C127 cells carrying BPV-1 mutant DNAs, identified by their ability to form foci, were picked and expanded into cell lines. For all of the mutants presented in this work, the BPV-1 DNA was found to be integrated into the cellular genome by isolating total cellular DNA and subjecting it to Southern blotting techniques (see the report of Berg et al. [3] for procedures). Genomic DNA from cell lines carrying mutant BPV-1 plasmids was digested with *Bam*HI and *Hind*III and analyzed by blotting to show that a BPV-1 fragment of 5.4 kilobases was present in these cells and therefore that the DNA sequences spanning the E1 frame and upstream regulatory region were intact. C127 cell lines carrying either pMT21 or pMT-E1 were made by cotransfecting these plasmids with pNEO5' at a ratio of 10:1 (pMT:pNEO5') and picking and expanding Neo^r colonies by the protocol described by Lusky and Botchan (33). These Neo^r colonies were screened for their ability to make the E1 protein by radiolabeling the cells and immunoprecipitating from the extracts with the α -M antiserum (see below).

pMT-E1, pMT-M, and pMT-Repy plasmids were transfected into COS cells by the DEAE-dextran procedure described by Kaufman et al. (20) with the following changes (R. Kaufman, personal communication). Following incubation with the DNA for 12 h, the cells in 10-cm-diameter dishes were shocked with 2 ml of 10% dimethyl sulfoxide reagent (137 mM NaCl, 5 mM KCl, 0.7 mM Na₂HPO₄, 6 mM D-glucose, 21 mM HEPES [N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid], 10% dimethyl sulfoxide [pH 7.1]) for 3 min at room temperature. Five milliliters of complete medium containing 0.1 mM chloroquine (Sigma Chemical Co.) was added per 10-cm-diameter plate for 2.5 h, after which the plates were rinsed once with phosphate-buffered saline (PBS), and complete medium was added. Cells were incubated for 48 to 72 h following chloroquine treatment before being radiolabeled for immunoprecipitation, prepared for immunostaining, or used to generate extracts for ATP affinity labeling.

Southern analysis of genomic DNA. Total genomic DNA was purified by standard modifications of the method of Thomas et al. (56). Southern blotting was done as previously described (3, 46).

Preparation of antisera. Anti-R polyclonal antisera were raised in New Zealand White rabbits against two bacterial fusion proteins essentially as described by Sigel et al. (51). Both fusion proteins contained the same 461-amino-acid segment encoded by the 3' portion of the E1 ORF (nt 1280 to 2663). Primary injections and initial boosts were done with a fusion protein expressed from the bacteriophage λ pL promoter in vector pAJ π (30) which consisted of the 3' E1 protein segment fused to 20 amino acids of the λ cII protein. The second boost was done with a fusion protein expressed from the bacteriophage T7 ϕ 10 promoter in vector pET-3c (43) which consisted of the 3' E1 protein segment fused to 13 amino acids of the T7 gene 10 protein. Fusion proteins were purified on preparative sodium dodecyl sulfate (SDS)-polyacrylamide gels and concentrated as described previously (57) before being injected.

Anti-M polyclonal antiserum raised against a 204-amino-

acid segment from the 5' portion of the E1 ORF (nt 860 to 1471) fused to the *Escherichia coli* TrpE protein has been described previously (57).

Radiolabeling procedures and immunoprecipitations. Proteins were labeled with ³²P or ³⁵S according to the protocols described by Thorner et al. (57) with the following modifications. Cells were starved for between 30 and 60 min and labeled for between 1 and 2 h (a 4-h period was used to label proteins with [³⁵S]methionine in COS cells). Extracts were prepared by washing the cell monolayer once with ice-cold PBS and adding 1 ml of ice-cold RIPA buffer (1–1 mM phenylmethylsulfonyl fluoride–1 μ M leupeptin–1 μ M pepstatin. The plate was swirled to cover the cell monolayer with the RIPA buffer, and after a few minutes, the cell extract was scraped into a microcentrifuge tube and kept on ice until all samples had been harvested. The samples were then centrifuged for 15 min at 4°C to compact the viscous DNA which was removed from the tube, and the cleared supernatant was used immediately for immunoprecipitation or stored at –70°C. TranS-label and [³⁵S]cysteine were used to label methionines and cysteines, respectively, and ³²PO₄ was used to label phosphates (ICN Pharmaceuticals Inc.).

Proteins were immunoprecipitated from radiolabeled extracts by either a one-step or a two-step immunoadsorption procedure (41) with the modifications described in reference 57 and below. The one-step procedure is the same as the first part of the two-step procedure, stopping at the point at which 30 μ l of 1% SDS in PBS is added to release the antigen and antibody from the immune complexes (57); instead, 2 $\times$ Laemmli sample buffer in preparation for SDS-polyacrylamide gel electrophoresis (PAGE) is added. The following changes were included in both the one- and two-step procedures. All incubations and manipulations were done on ice except when otherwise noted. RIPA buffer was substituted for solubilization buffer, and all solutions contained 1 mM phenylmethylsulfonyl fluoride–1 μ M leupeptin–1 μ M pepstatin. Lysates were precleared by incubating them with 50 μ l of nonspecific rabbit serum (Vector Laboratories) for 30 to 60 min and then with 100 μ l of a 50% suspension of Pansorbin (Calbiochem-Behring) for an additional 15 to 30 min. The samples were then centrifuged for 10 min, and the supernatants were transferred to a new tube, where they were incubated with specific rabbit serum as previously described. Antigen-antibody complexes were precipitated with either 25 μ l of Pansorbin (10% suspension) pretreated with 2% bovine serum albumin (BSA) or 200 μ l of protein A-Sepharose (10% suspension; Pharmacia) pretreated with 3% BSA. Immune complexes were washed twice with RIPA buffer, suspended in 1 ml of high-salt buffer (125 mM Tris hydrochloride [pH 7.5], 12.5 mM dithiothreitol, 6.25 mM CaCl₂, 0.5 M NaCl), incubated at 37°C for 30 min, and then washed once more in RIPA buffer and once in T₁₀E₅ (10 mM Tris hydrochloride [pH 7.5], 5 mM EDTA). For the two-step immunoprecipitations, the high-salt step was typically included in the washes only following the second round of antibody incubation. The immunoprecipitated proteins were removed from the final pellet and subjected to SDS-PAGE as described previously (57).

Phosphatase treatment of the E1 protein. Extracts from radiolabeled COS cells transfected with pMT-E1 were immunoprecipitated with the α -M antibody by the one-step procedure, and during the final wash, the T₁₀E₅ containing the immune complexes was split into three equal portions. These portions were centrifuged, the T₁₀E₅ was removed, and the pellets were suspended in 100 μ l of one of the following buffers and treated with the indicated phosphatase.

For calf intestinal alkaline phosphatase (CIP) treatment, 100 mM Tris hydrochloride (pH 9.0) and 1 U of the enzyme (Boehringer Mannheim Biochemicals) were used. For potato acid phosphatase (PAP) treatment, 100 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) (pH 6.0) and 0.5 U of the enzyme (Boehringer) were used according to the procedure of Buchkovich et al. (6). As a control, 100 mM MES (pH 6.0) and no enzyme were added to the third pellet. All reaction mixtures were incubated at 37°C for 30 min and then washed once with $T_{10}E_5$ and subjected to SDS-PAGE.

ATP affinity labeling of the E1 protein. Extracts from COS cells that had been transfected with pMT-E1 were prepared as follows. All steps after the first PBS wash were performed on ice and used ice-cold buffers. The cell monolayers were washed once with PBS and once with hypotonic buffer (20 mM Tris hydrochloride [pH 7.6], 5 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM dithiothreitol) scraped into microcentrifuge tubes, and pelleted. After the buffer was removed, the cells were resuspended in 250 μ l of hypotonic buffer and lysed by the addition of 250 μ l of hypotonic buffer–1% Triton X-100 (Sigma), and the nuclei were pelleted in a microcentrifuge for 5 min at 4°C. Soluble nuclear extracts were made by treating the nuclei with 50 μ l of hypotonic buffer–350 mM NaCl for 20 min on ice and then removing the nuclei by centrifuging for 10 min. Reactive [α - ^{32}P]2', 3'-dialdehyde ATP (ox-ATP) was made by a modification of the method described by Clertant and Cuzin (9). [α - ^{32}P]ATP (Amersham Corp.; 10 mCi/ml; >400 Ci/mmol; 80 μ Ci/12 μ l of reaction mixture) was oxidized in 4 mM HCl–4 mM $NaIO_4$ (Sigma) and incubated for 30 min in the dark. The reaction was stopped by the addition of 0.25 volume of 50% glycerol for an additional 20 min to reduce any unreacted periodate. This ox-ATP was then added to nuclear extracts in a final reaction mixture containing 0.5 M KCl–5 mM $MgCl_2$ –1 mM dithiothreitol–1 mM EGTA [ethylene glycol-bis(β -aminoethyl ether)-*N,N,N',N'*-tetraacetic acid] (100- μ l final reaction volume). After 30 min, 0.02 volume of 0.5 M $NaCNBH_4$ (Sigma) was added, and the cross-linking reaction mixture was incubated on ice overnight. The protein mixture was then immunoprecipitated with the α -M antibody according to the one-step procedure and subjected to SDS-PAGE.

Immunostaining. COS cells were trypsinized 2 days after transfection, plated onto glass microscope slides placed in square 10-cm dishes, and allowed to adhere to the slides overnight. The slides were washed once in PBS, fixed in an acetic acid-ethanol solution (1:20) at –20°C for 30 min, permeabilized with PBST (PBS, 0.1% Triton X-100) for 10 min at room temperature, and then blocked with 1% BSA in PBST for 15 min. Primary antibody (α -M for detecting M and E1 or α -R for detecting E1 and Rep γ), diluted 1:150 in PBST–1% BSA, was added to the slides, and the slides were incubated for 2 h in a wet chamber. The slides were then washed four times with PBST–0.1% BSA for 5 min each time with rocking. Horseradish peroxidase-conjugated goat anti-rabbit antibody (Vector), diluted 1:2,000 in PBST–1% BSA, was then added to the slides. These were incubated and washed as described above. The slides were developed with a solution of PBS–0.5 mg of diaminobenzidine (Polyscience Inc.) per ml–0.03% H_2O_2 (Sigma)–0.04% $NiCl_2$. The development reaction was stopped after 10 to 30 min by washing with PBS.

RESULTS

Overexpression and detection of a 68-kDa phosphoprotein encoded by the entire E1 ORF in COS cells. To define the

conditions necessary to identify a putative E1 protein, we first overproduced this protein in COS cells using the transient expression plasmid pMT-E1 (Fig. 1A). In addition to using extracts greatly enriched for the E1 protein (because of the amplification of the signal afforded by the vector), we also modified our immunoprecipitation protocol from that used previously (57) to permit clearer visualization of proteins in the high-molecular-weight range. Figure 1B shows that the α -M antibody recognizes and can immunoprecipitate two proteins of apparent molecular masses of 68 and 75 kDa in extracts metabolically labeled with ^{35}S from COS cells transfected with pMT-E1. These proteins were present only in extracts from pMT-E1-transfected cells and not from cells transfected with the vector alone (Fig. 1B, lanes 1 and 2). In addition, these two proteins were also visible on Western immunoblots with α -M antibody (done as described by Thorner et al. [57]) in extracts of COS cells transfected with pMT-E1 but not in those transfected with pMT21 (data not shown). The additional smaller proteins in lane 2 appear to be related to the E1 protein, since they were present only in extracts from pMT-E1-transfected cells, but because they varied in size and intensity from experiment to experiment (data not shown), they probably represent proteolytic degradation products.

Because the M protein is known to be phosphorylated (57), it seemed likely that a full-length E1 protein would also be phosphorylated. Indeed, immunoprecipitation and SDS-PAGE analysis of ^{32}P -labeled extracts of cells transfected with pMT-E1 revealed the presence of two phosphoproteins having the same apparent molecular masses as the major ^{35}S -labeled proteins (Fig. 1B, lanes 2 and 4). These two proteins were the only phosphoproteins in pMT-E1-transfected cells recognized by the α -M antibody and were absent from pMT21-transfected cells (Fig. 1B, lanes 3 and 4). The protein of 68 kDa is of a size consistent with the calculated molecular mass of a protein initiating at the first ATG in E1 and translating through to the stop codon at the end of the ORF. It may be that the larger protein results from some form of posttranslational modification of the 68-kDa species. This idea is discussed in more detail below.

Identification of a bona fide full-length E1 protein in BPV-1-infected cells. To determine whether a large E1 protein actually exists in BPV-1-transformed cells, we applied the same immunoprecipitation conditions used to identify E1 in COS cells to radiolabeled extracts from BPV-1-infected C127 cells (VI216 cells). In addition, we analyzed extracts from C127 cells that contained integrated copies of the plasmids described in Materials and Methods. As shown in Fig. 2A, two proteins were immunoprecipitated by the α -M antiserum from BPV-1-positive (VI216) cells labeled with ^{32}P (lane 8) but were absent in the BPV-1-negative cells (either C127 cells alone [Fig. 2C, lanes 1 and 5] or C127 cells containing pMT21 [Fig. 2A, lane 1]). In addition to the 23-kDa M protein previously identified in BPV-1-transformed cells, a protein with an apparent molecular mass of 68 kDa was recognized by the α -M antiserum in VI216 cells. This 68-kDa protein comigrated with the full-length E1 protein produced in C127 cell lines containing pMT-E1 (Fig. 2A, lanes 2 to 4 and 6), making it a good candidate for the product of the E1 ORF. This 68-kDa protein was also visible in immunoprecipitates from ^{35}S -labeled VI216 cells (Fig. 2B, lane 3) but not from C127 cells (lane 4). As in the experiments described above, in these immunoprecipitation experiments using ^{35}S -labeled extracts the 68-kDa protein in VI216 cells comigrated with the E1 protein produced from the pMT-E1 vector in C127 cells (Fig. 2B, lanes 1 to 3),

FIG. 1. (A) Organization of the transforming region of BPV-1 DNA and E1 ORF proteins produced from the pMT-E1, pMT-M, and pMT-Repy expression vectors. The top line shows the physical map of the transforming fragment of BPV-1 which extends from the *Hind*III site at nt 6959 to the *Bam*HI site at nt 4451. Open boxes numbered from E1 to E8 above the map represent the eight ORFs in the BPV-1 early region and are aligned according to reading frame. Five promoters (P1 to P5) found in this region are also shown. The *M* and *R* genes as defined genetically (3, 35) are shown as double lines below the map. Open boxes below the map show the structures of proteins encoded by the E1, Rep γ , and *M* (spliced from nt 1235 to nt 3225) fragments of the E1 ORF (nt 813 to nt 2663) expressed from the adenovirus major late promoter (Ad MLP) in pMT 21 (shown at the left). In addition to the adenovirus major late promoter, pMT21 contains the adenovirus tripartite leader (TPL), an intervening sequence (IVS), the dihydrofolate reductase gene (*dhfr*), the SV40 polyadenylation site (polyA), the adenovirus VA₁ gene, the pBR322 origin (*ori*), an ampicillin resistance gene (*AmpR*), and the SV40 origin (SV40 *ori*) and enhancer. (B) Autoradiogram showing the E1 proteins (arrows) after immunoprecipitation from pMT-E1 (E1)- and pMT21 (V)-transfected COS cells labeled with either ³⁵S (lanes 1 and 2) or ³²P (lanes 3 and 4). Immunoprecipitations were done with α -M serum by the one-step immunoadsorption procedure. kd, Kilodaltons.

suggesting that these proteins are the same. The 68-kDa protein detected in virus-infected C127 cells also comigrated with the 68-kDa form produced in COS cells (data not shown). The 68-kDa protein was more readily detected by ³²P labeling than by ³⁵S labeling. The Rep γ protein (Fig. 1) was as heavily labeled by ³²P as was the 68-kDa protein when produced in COS cells (data not shown). Thus, we used the phosphate label as a positive tag for the 68-kDa protein in the experiments described below.

To see if this protein is indeed encoded by the continuous E1 ORF sequence, we performed a number of experiments. In the first of those experiments, ³²P-labeled extracts from BPV-1-infected and uninfected cells and cells containing the BPV-1 mutant plasmid i2113-2 were immunoprecipitated according to a two-step immunoadsorption procedure wherein two different antisera were used in succession. All of the extracts were first immunoprecipitated with the α -M antiserum. Immune complexes were then dissociated by the addition of PBS containing 1% SDS and by being boiled for

5 min to disrupt all antigen-antibody interactions and to irreversibly destroy the antibodies. After the SDS was diluted to <0.16%, a second antiserum was added (either α -M serum, α -R serum, or nonimmune serum). Figure 2C shows that the E1 protein found in either VI216 or ID13 cells immunoprecipitated in the first round by a α -M can be immunoprecipitated by either α -M or α -R in the second round (lanes 3, 4, 7, and 8). However, since the α -R antibodies were raised against a 3' portion of the E1 ORF fused to bacterial sequences, the *M* protein was immunoprecipitated only when α -M serum was used for the second round (lanes 3 and 4). This is because the E1 determinants which α -R serum recognizes fall outside the *M* domain. In addition, neither protein was immunoprecipitated when the second round was done with nonimmune serum (lanes 9 to 12), indicating that the α -M antibodies from the first round were indeed rendered nonfunctional and that the appearance of E1 and *M* following SDS-PAGE depended on the specificity of the second antiserum. The i2113-2 mutation lies in

FIG. 2. A 68-kDa protein encoded from the E1 ORF is present in BPV-1-infected cells. (A) Autoradiogram showing that a protein with the same apparent molecular mass as a protein expressed from the whole E1 ORF is present in BPV-1-infected cells (VI216). Proteins were immunoprecipitated from 32 P-labeled C127 cells containing pMT-E1 or pMT21 plasmids and VI216 cells with α -M serum by the two-step immunoadsorption procedure. The upper arrow points to the E1 protein present in four of the C127-E1 cell lines (lanes 2 to 4 and 6) and VI216 cells (lane 8) but not in two of the C127-E1 cell lines (lanes 5 and 7) or the C127-pMT21 cell line (lane 1). The two C127-E1 cell lines showing no E1 protein (lanes 5 and 7) may have received only the pNEO5' plasmid during transfection, or the pMT-E1 plasmid may exist in these cells integrated in a way that prevents E1 synthesis. The lower arrow points to the M protein, which is present only in VI216 cells (lane 8). (B) Autoradiogram showing that the E1 protein immunoprecipitated from 35 S-labeled VI216 cells (lane 3) comigrates with the E1 protein immunoprecipitated from 35 S-labeled C127 cells containing pMT-E1 (lanes 1 and 2). The M protein is present only in VI216 cells (lane 3). Neither protein is present in C127 cells (lane 4). (C) Autoradiogram showing proteins immunoprecipitated from 32 P-labeled C127 cells, i2113-2 cells, VI216 cells, and ID13 cells by the two-step immunoadsorption procedure. Proteins from all extracts were first immunoprecipitated with α -M serum, and then a second immunoprecipitation was

performed with either α -R (lanes 5 to 8), or nonimmune (lanes 9 to 12) serum. The upper arrowhead points to the E1 protein immunoprecipitated from VI216 (lanes 3 and 7) and ID13 (lanes 4 and 8) cell extracts, but not from C127 (lanes 1 and 5) or i2113-2 (lanes 2 and 6) extracts, by both α -M and α -R sera. The lower arrowhead points to the M protein immunoprecipitated by α -M serum from i2113-2 (lane 2), VI216 (lane 3), and ID13 (lane 4) extracts but not from C127 (lane 1) extracts. Neither protein was immunoprecipitated from any of the extracts when nonimmune serum was used for the second immunoprecipitation (lanes 9 to 12). (D) The 68-kDa protein can be immunoprecipitated by the α -R serum alone. The E1 protein is not present in the C127 control (lane 1) but is evident in two different BPV-infected cell lines (lanes 2 to 4). Lanes 2 and 3 are immunoprecipitations from cells containing the BPV genome with a mutation in the ATG codon at nt 3093 (55). Lane 2 is from unsynchronized cells. Lane 3 is from a culture that was synchronized by serum starvation and labeled in the S phase, which occurred 16 h after release. Lane 4 is from an unsynchronized VI216 culture (wild-type BPV). The 68-kDa E1 protein is the only phosphoprotein in BPV-infected cells recognized by the α -R serum that is not present in the uninfected C127 parental line. Lane 5 shows a VI216 cell extract immunoprecipitated by the α -M serum in which both the 68-kDa E1 and M proteins are recognized. All extracts were labeled with 32 P for 2 h.

the 3' portion of the E1 ORF, and BPV-1 plasmids with this mutation have been shown previously to make M (57). Figure 2C confirms this result and shows in addition that the large E1 protein is absent in the i2113-2 mutant (lanes 2 and 6; see also Fig. 3C and results described below). In Fig. 2D, we show that the anti-R serum recognizes the 68-kDa protein when it alone is used as the immunoprecipitating reagent. The epitopes recognized by the anti-R sera are encoded by the 3' portion of the E1 ORF. We conclude that if other E1-encoded phosphoproteins are produced in these cells, they fall below our level of detection.

Detection of the 68-kDa phosphoprotein correlates with the genetics of the E1 ORF. To examine more rigorously the structure of the gene encoding the 68-kDa protein, we analyzed protein expressed from a collection of translation-termination mutations placed throughout the E1 ORF (Fig. 3A). In addition, a mutation (1M-I) that changes the first methionine codon (ATG) in the E1 ORF to an isoleucine codon (ATA) was studied. In agreement with our previous studies, we found that mutations which inactivated 5' E1 functions had significantly lower transformation efficiencies. All of these mutations abolished plasmid replication, as the BPV-1 DNA present in the foci that were obtained and expanded as described in Materials and Methods was always found integrated into the host cell genome (see Fig. 5C; additional data not shown). Southern blot analysis of genomic DNAs showed that the BPV-1 DNA remained intact and existed in multiple copies in all cell lines. This finding is consistent with, but does not prove, the idea that the BPV-1 proteins detected and described below are not likely to be the result of new transcription units created by integration (data not shown). Figure 3A shows the sizes expected for the M and E1 proteins in the different mutants, assuming that M and E1 have the structures anticipated by the wild-type M and wild-type E1 diagrams at the bottom of Fig. 3A. Immunoprecipitation of 32 P-labeled extracts of cell lines containing various mutant E1 ORFs yielded results consistent with the 68-kDa protein being encoded by sequences colinear with the ORF (Fig. 3B through D). Moreover, the 68-kDa protein was absent in all 1M-I cell lines examined (Fig. 3B, lanes 5 to 12), suggesting that the methionine codon at nt 849 is in fact the site of translational initiation of the E1 protein.

FIG. 3. Detection of the 68-kDa phosphoprotein correlates with the genetics of the E1 ORF. (A) The BPV-1 E1 ORF (nt 813 to 2663), showing the locations of the various mutations used to explore the gene structure of the 68-kDa E1 protein. The nucleotide position of each mutation is given immediately above the E1 ORF. The E1-encoded protein products that the mutants are expected to produce are shown as small boxes. Mutants with lesions before the splice donor at nt 1235 are expected to make only one E1 product (i.e., E1Sma, 118oc, and 128am). Mutants with lesions after the splice donor are expected to make two E1 products, M and a truncated E1 protein (i.e., 131am and i2113-2). The shaded regions represent the 13-amino-acid exon spliced from nt 3225 (8, 57). The presumed structures of the wild-type (wt) M and E1 proteins are shown at the bottom of the figure. (B through D) Autoradiograms showing E1 and E1-related proteins immunoprecipitated with α -M serum from 32 P-labeled cell lines harboring the mutant BPV-1 DNAs described in panel A by the one-step immunoadsorption procedure. (B) The arrows point to the E1 and M proteins present in the VI216 (lane 2) cell line but not in the C127 (lane 1), E1Sma (lanes 3 and 4), or 1M-I (lanes 5 to 12) cell line. (C) The upper arrow points to the E1 protein present in the VI216 (lane 12) cell line but not in the C127 (lane 1) or i2113-2 (lanes 2 to 11) cell lines. The lower arrow points to the M protein present in the VI216 (lane 12) and i2113-2 (lanes 2 to 11) cell lines but not in the C127 (lane 1) cell line. (D) The arrow points to the E1 protein present only in the VI216 (lane 2) and 434P-S (lane 7) cell lines. Smaller E1 proteins (bracket) are produced in the 118oc (lane 3), 128am (lane 4), and 131am (lane 5) cell lines. Also included in the bracket is the M protein produced in VI216 (lane 2), 131am (upper band of doublet in lane 5), i2113 (lane 6), and 434P-S (lane 7). No E1 or E1-related proteins are produced in C127 cells (lane 1). kd, Kilodaltons.

It otherwise seemed unlikely that this reasonably conservative amino acid substitution (17) should eliminate cross-reacting E1 protein. Interestingly, although the 1M-I mutation greatly reduced the amount of M present in these cells (by a factor of about 100), some M protein was visible for all these cell lines upon longer exposure of the film (data not shown). This finding reinforces the notion that the nt 849 Met codon starts both proteins, as it has previously been shown that translation can initiate at codons other than ATG, particularly the ATA codon, but with greatly reduced efficiencies (40).

No 68-kDa E1 or M protein was observed in the cell lines containing the E1Sma mutant (Fig. 3B, lanes 3 and 4). This is consistent with the interpretation that these mutants synthesize a small E1 polypeptide of 4.8 kDa (Fig. 3A) which either runs off of the gel under our SDS-PAGE conditions or is unstable. Similarly, although no large 68-kDa protein could be detected in cell lines harboring the 3' E1 BPV-1 mutant i2113-2, a 23-kDa (M) protein was detected (Fig. 3C, lanes 2 to 11). The results obtained with the termination mutants placed either 5' or 3' to the splice donor at BPV-1 nt 1236 provide evidence for the co-amino-terminal structures

FIG. 4. Phosphatase treatment of ^{32}P - and ^{35}S -labeled E1. The autoradiogram shows E1 (arrows) that was first immunoprecipitated from pMT-E1-transfected COS cells labeled with ^{32}P (lanes 1 to 3) or ^{35}S (lanes 4 to 6) and then treated with either CIP (lanes 2 and 5), PAP (lanes 3 and 6), or no phosphatase (lanes 1 and 4). The one-step immunoadsorption procedure was performed with α -M serum.

of the 68- and 23-kDa proteins. Truncated proteins with predicted patterns were produced by these mutants (Fig. 3D). Our ability to detect protein fragments of the 68-kDa protein whose sizes were similar to that of the first exon of the M protein, but not significantly larger fragments (e.g., those produced by the i2113-2 mutants), may be ascribed to the instability of these latter proteins. Mutant 118oc, which contains a termination codon at nt 1200, produced neither an M protein nor a 68-kDa protein, but showed a novel protein band migrating below M (Fig. 3D, lane 3). Mutant 128am, which contains a termination codon at nt 1230, produced another novel E1 cross-reacting fragment migrating faster than M but slower than the one produced by 118oc (Fig. 3D, lane 4). Similarly, mutant 131am, which contains a termination codon just after the splice donor, produced a wild-type M protein with a novel E1 fragment protein comigrating with the fragment detected with the 128am mutant. We point out that the structural difference predicted for the E1 fragments produced by 128am and 131am is only three amino acids. Therefore, the electrophoretic mobility differences would likely be undetectable in the gel system used. As we showed previously (57), the *in vitro* translation product of the M gene has an anomalously low electrophoretic mobility. Similarly, all of the novel E1 fragments described here also migrated slowly. However, the relative difference between them is almost precisely what is predicted by the diagram in Fig. 3A.

Phosphatase treatment of the E1 protein. As shown in Fig. 1B (lanes 2 and 4), expression from pMT-E1 in COS cells generated two forms of the E1 protein, both of which are phosphorylated. In contrast, BPV-1-infected cells showed only one form of the E1 protein, also phosphorylated (Fig. 2A, lane 8, and Fig. 2B, lane 3). To ask if the two forms in COS cells were generated by differential phosphorylation of a single polypeptide, we performed the following experiment. Extracts of COS cells labeled with either ^{32}P or ^{35}S were immunoprecipitated with α -M serum and the antigen-antibody-protein A-Sepharose pellets were then treated with either CIP or PAP prior to SDS-PAGE. As shown in Fig. 4, treatment with CIP removed many of the phosphates from the two proteins (lane 2), whereas treatment with PAP removed them all (lane 3). However, both protein species persisted in the same relative abundance after treatment with

PAP as described before (Fig. 4, lanes 4 to 6). Thus, we suspect that the difference in mobility between the two forms is due to some other type of modification. It is interesting that the two forms of the E1 protein exhibited similar responses to phosphatase treatment. CIP removed the same proportion of phosphates from each of the two E1 proteins, while PAP removed all of the phosphates from both proteins (Fig. 4, lanes 1 to 3). In addition, although this is not obvious from Fig. 4, which is overexposed, these two proteins were present at the same ratio in cells labeled with ^{32}P as in cells labeled with ^{35}S before and after phosphatase treatment (lanes 1, 2, and 4 to 6 and data not shown). Taken together, these results suggest that the two forms of E1 in COS cells contain the same number of phosphates. The protein band migrating at 75 kDa was not seen in C127 cells stably transformed with the expression vector pMT-E1, and the pattern of proteins seen in COS cells was not affected by mutating the donor at nt 1236 (data not shown). However, this 75-kDa protein, which was less abundant than the 68-kDa protein (Fig. 1B and 4), may be present in the stably transformed cell types, but in amounts below our limit of detection. Possibly E1 is more susceptible to posttranslational modification in COS cells than in C127 cells, either because of the unnaturally high amounts of the protein produced in COS cells or because of some other feature of the COS cell environment.

The E1 protein binds ATP. Comparison of protein-coding sequences has shown that a carboxy-terminal portion of the BPV-1 E1 protein shares a degree of similarity with the known ATP-binding domains of polyomavirus and SV40 large-T antigens (10). We performed affinity-labeling experiments to determine whether the E1 protein produced in COS cells can be radioactively labeled by the oxidized ATP derivative ox-ATP. This ATP analog couples covalently to primary amines on ATP-binding proteins. The ox-ATP reactions were performed with unlabeled nuclear extracts, and then α -M serum was used to immunoprecipitate the E1 protein for subsequent analysis by SDS-PAGE and autoradiography. In parallel, another batch of the same cells used to prepare extracts was labeled with ^{35}S and immunoprecipitated with α -M serum to monitor E1 synthesis. Immunoprecipitation of E1-containing extracts treated with ox-ATP showed that E1 did indeed bind ATP as predicted. Furthermore, both forms of the E1 protein were efficiently labeled and hence must have bound ATP (Fig. 5A, lane 5). As a control, nuclear extracts from pMT21-transfected COS cells which do not make E1 (lane 1) were also treated with ox-ATP and immunoprecipitated. No labeled proteins were detected (lane 4). To ask specifically if the binding of ox-ATP results from interaction with the putative ATP pocket predicted by sequence inspection, we constructed a mutation in this domain and performed an analogous experiment with the appropriate expression vector. The mutation 434P-S, which changes the proline at amino acid position 434 in E1 to a serine (Fig. 3A and 5), was selected because of its analogy to the SV40 T-antigen mutant, tsA209 (12, 49). The mutation in the SV40 genome resides in those sequences encoding the putative Rossmann mononucleotide-binding fold (4, 44, 45) and renders the virus temperature sensitive for DNA replication. Southern blotting analysis showed that BPV-1 plasmids carrying the analogous mutation 434P-S could not replicate at either 31 or 37°C and were instead found integrated into the host cell genome (Fig. 5B). In these lines the integrated viral DNA was still able to synthesize both the E1 and M proteins (Fig. 3D, lane 7). When pMT-E1 plasmids carrying this mutation were transfected into COS cells, E1

FIG. 5. (A) E1 produced in COS cells can be radiolabeled with periodate-oxidized [^{32}P]ATP (ox-ATP). Extracts prepared from COS cells transfected with either pMT21 (V), pMT-E1 (E1), or pMT-E1 containing the 434P-S mutation (434P-S) were treated with ox-ATP and then immunoprecipitated (lanes 4 to 6). In parallel, COS cells transfected with the same plasmids (V, E1, and 434P-S) were labeled with [^{35}S]methionine and the extracts were immunoprecipitated (lanes 1 to 3). Immunoprecipitations were done with α -M serum according to the one-step immunoabsorption procedure. The E1 protein (arrows) produced in the E1-transfected cells (lane 2) is radiolabeled by ox-ATP (lane 5). The E1 protein produced in the 434P-S-transfected cells (lane 3), however, is not radiolabeled by ox-ATP (lane 6). No E1 is produced in the cells transfected with the pMT21 plasmid (lanes 1 and 4). The homologous amino acid sequences for BPV-1 E1 and the SV40 large-T antigen are shown at the bottom, with the proline-to-serine change that gives rise to the 434P-S mutation (BPV-1) and the analogous *tsA209* mutation (SV40) indicated. (B) 434P-S is defective for DNA replication at both 31 and 37°C. Genomic DNA from C127 cells stably transformed by BPV-1 434P-S (lanes 1 to 8) or wild-type BPV-1 (VI216; lane 9) was digested with *Xho*I, which does not cut within the BPV-1 plasmid. In all 434P-S lines examined, whether at 31°C (lanes 1 to 4) or at 37°C (lanes 5 to 8), no free BPV-1 DNA is detected. (C) 1M-I does not replicate in stably transformed cell lines. Uncut total DNA (10 μg per lane) from three independent BPV 1M-I cell lines (lanes 1 to 3)

protein was made (Fig. 5A, lane 3); however, when these extracts were treated with ox-ATP, this protein was not labeled, indicating that it was in fact defective for ATP binding (Fig. 5A, lane 6). From these results, we conclude that E1 does indeed bind ATP in a specific manner and that this binding is probably necessary for BPV-1 plasmid replication. It is curious that the 434P-S mutant produced only one form of the E1 protein (68 kDa) in COS cells (Fig. 5A, lane 3) at consistently lower quantities. It is tempting to speculate that the absence of ATP binding precludes a conformational change in E1 that is necessary for some unknown modification to occur. This modification could then result in a slower gel mobility. Similarly, ATP binding may be needed to stabilize E1 and accumulate the higher levels normally observed in COS cells for the wild-type protein. The 68-kDa form could also be a degradation product of the larger form. However, we do not favor this hypothesis, as the two bands in COS cell extracts were always present in about the same ratios regardless of the storage time of the extracts.

The E1 protein is localized in the cell nucleus. Because the products of the E1 ORF have been implicated in BPV-1 plasmid replication, we were interested in knowing whether the full-length E1 protein was localized in the nucleus. COS cells transfected with pMT-E1 were used to determine the subcellular localization of the E1 protein. In addition, we transfected COS cells with the expression vector encoding M (pMT-M), the expression vector encoding a fragment of E1 (pMT-Repy), and the vector alone (pMT21). Intracellular distribution was visualized by immunostaining with either α -M or α -R serum, and approximately 10% of the cells showed staining, indicating that they contained constructs producing cross-reacting antigens. All the stained cells in the population transfected with pMT-E1 that possessed clearly visible nuclei exhibited almost exclusively nuclear staining with both the α -M and α -R sera (Fig. 6C and D). Cells transfected with the pMT21 vector without an insert gave no discernable signal with either α -M or α -R serum (Fig. 6A and B), demonstrating that the signal was indeed specific to E1 and related proteins. Furthermore, the M protein made by pMT-M and visualized by staining with the α -M serum was also localized solely to the nucleus (Fig. 6E). In striking contrast, the Repy protein made from pMT-Repy and recognized with the α -R serum was found only in the cytoplasm (Fig. 6F). From these results, we conclude that the full-length E1 and M proteins reside in the nucleus of the cell, presumably because they contain a nuclear localization sequence that is missing in Repy. Interestingly, Clertant and Seif (10), in their sequence analysis of the E1 ORF sequence, noticed a second region of homology between the papillomavirus and SV40 and polyomavirus large-T antigens. This region, which they called X, is in the 5' end of the BPV-1 E1 ORF. Within this region of homology in the large-T antigens reside the known nuclear localization signals (11, 18, 19, 28, 29, 42, 53). We point out that in the 5' amino-terminal domain of E1, at amino acid positions 103 to 109 (counting the first Met as 1), there exists a strong homology to the basic residues in large-T antigens known to be critical for nuclear localization. We are currently testing the importance of these amino acids for nuclear localization of E1.

and VI216 cells (lane 4) were analyzed by Southern blotting. No free BPV is detectable in the 1M-I cell lines. Nick-translated BPV DNA was used as the probe in panels B and C. Numbers above the gels indicate the particular cell line in each lane.

FIG. 6. The E1 and M proteins reside in the nucleus of the cell. Shown are phase-contrast photomicrographs taken after immunostaining of COS cells transfected with pMT21 (A and B), pMT-E1 (C and D), pMT-M (E), or pMT-Repy (F). Cells attached to slides were incubated with either α -M serum (anti-M) or α -R serum (anti-R). The specific antibodies were detected with horseradish peroxidase-conjugated goat anti-rabbit antibody and developed as described in the text.

DISCUSSION

In this report we present evidence that a 68-kDa protein encoded by BPV-1 is synthesized from the viral E1 ORF. Translation of the protein initiates from the first methionine codon in the ORF, and the E1 gene appears to be colinear with the entire ORF. The antiserum used to identify the full-length E1 protein is the same antiserum that was used to identify the product of the *M* gene in an earlier report (57). In that study, analysis of higher-molecular-mass proteins (>46 kDa) immunoprecipitated from ^{35}S - and ^{32}P -labeled cells was hampered by nonspecific precipitation of high-molecular-mass background bands on the autoradiographs; thus, any band at 68 kDa would have been completely obscured. When

the background levels were low in this region of the gel, the total signal was low, and thus we missed detection of the 68-kDa protein. Optimizing conditions for detecting the E1 protein expressed in COS cells enabled us to develop an immunoprecipitation procedure which eliminated the high background level seen previously. This development, in turn, led to our discovery of a 68-kDa phosphoprotein present in BPV-1-infected cells. The changes in procedure which facilitated our identification of the full-length E1 product included (i) using RIPA buffer containing 0.1% SDS to solubilize the labeled cells before immunoprecipitation instead of the solubilization buffer used previously (57), which contained only nonionic detergents; (ii) removing the

DNA from the cell extracts before immunoprecipitation rather than shearing it; (iii) clearing the cell extracts with large amounts of nonspecific rabbit serum and Pansorbin; and (iv) washing the immune complexes with a high-salt-containing buffer at 37°C for 30 min before subjecting them to SDS-PAGE.

Although we believe that this report is the first clear identification of the 68-kDa papillomavirus E1 protein, it is consistent with two earlier findings. Seedorf et al. (48) could not identify an E1 protein (or E1-related products) in human papillomavirus type 16 (CaSki and SiHa)- or type 18 (HeLa, C4-1, and SW756)-transformed cells. However, their E1 serum could identify a 70-kDa protein translated *in vitro* in a rabbit reticulocyte system primed by hybrid selected RNAs. While it is possible that this translation product was simply produced by artificial starts from nuclear precursor RNAs or truncated RNAs, it is likely to represent a bona fide human papillomavirus counterpart of the 68-kDa protein described here. In addition, in a note added in proof to a recent publication, Burnett et al. (7) mentioned having detected an E1 antigen recognized by antisera directed against both 5' and 3' E1 ORF determinants.

The E1 ORF clearly encodes two proteins: the 68-kDa and M proteins. The 23-kDa protein cannot be considered a degradation product of the 68-kDa protein. Mutations which inactivate the splice donor at BPV-1 nt 1236 abolish M protein production but still allow for full-length 68-kDa protein, and the expression vector pMT-E1 produces only 68-kDa protein. In both instances, one would expect an apparent 23-kDa protein if it were simply a degradation product of the 68-kDa protein. Furthermore, mRNAs capable of encoding the 23-kDa protein have been identified by cDNA cloning (8), and translation *in vitro* of transcripts of these cDNAs yields a product that comigrates with *in vivo*-labeled protein (57).

The full-length E1 protein possesses features consistent with its playing a role in BPV-1 replication. In keeping with its high degree of similarity to the ATP-binding domains of the T antigens of SV40 and polyomavirus, the full-length E1 protein has been shown to bind ATP. A point mutation, 434P-S, in the carboxy terminus of E1 abolishes this activity, thereby establishing the specificity of the ATP-binding activity seen with wild-type E1. In addition, BPV-1 plasmids carrying the 434P-S mutation are defective for DNA replication in stable transformation assays, suggesting that the ATP-binding activity is required for E1 function in BPV-1 replication. P. Clertant, S. Santucci, R. Lathe, M.-P. Kieny, and E. J. Androphy have observed affinity-labeling results in agreement with our findings (P. Clertant, personal communication). Because the SV40 T antigen has been shown to possess an ATPase activity which is required for its function as a helicase in viral DNA replication (13, 54), it is interesting to speculate that the full-length E1 protein may play an analogous role in BPV-1 replication, and we are currently testing this idea. The finding that the E1 protein resides in the cell nucleus is also consistent with its putative involvement in viral DNA replication. A possible nuclear localization signal exists in the 5' portion of the E1 ORF at nt 1155 to 1175. This putative signal sequence in E1 (Pro-Val-Lys-Arg-Arg-Lys-Ser) is very similar to that of the SV40 T antigen (Pro-Lys-Lys-Lys-Arg-Lys-Val [19]) and is in that part of the ORF which encodes M as well as the 68-kDa protein. As we have shown here, the M protein is also a nuclear protein. Because previous transient-replication and stable-complementation experiments suggested that the 3' portion of E1 could encode a protein which was functional in

BPV-1 replication (see below), we were interested in testing whether a candidate for this protein would localize to the nucleus. We found that when the Repy protein encoded by the 3' portion of E1 was expressed in COS cells, this protein was, in fact, not localized in the nucleus. Consequently, if a carboxy-terminal fragment of E1 is to get into the nucleus, it must "piggyback" on another nucleoprotein or be spliced to an additional exon(s) encoding a nuclear localization signal.

The full-length E1 protein and the M protein are phosphorylated. We report here that the phosphates in the 68-kDa protein can be partially removed by treatment with CIP and completely removed by treatment with PAP. Interestingly, the SV40 T antigen exhibits this behavior with regard to these two phosphatases, and in the case of T antigen it is known that CIP dephosphorylates phosphoserine but not phosphothreonine (50), whereas PAP dephosphorylates both (22). The phosphorylation state of these amino acids has both positive and negative effects on the ability of T antigen to bind the SV40 origin and replicate DNA (16, 38, 52). For example, unphosphorylated T antigen cannot promote DNA replication *in vitro* but can be activated by phosphorylation of the threonine at amino acid position 124 (36, 37). Recent work has shown that this threonine can be phosphorylated *in vitro* by the cdc2 protein kinase, an important regulator of the cell cycle in eucaryotes (36, 39). BPV-1 E1 contains a threonine at the same relative position with respect to its postulated nuclear localization signal as does the SV40 large-T antigen. To press the analogy further with the SV40 large-T antigen, it will be important to know if the threonine 102 in E1 can be phosphorylated by cdc2 *in vitro* and what effect mutation of this residue has on BPV-1 replication. SV40, polyomavirus, and papillomaviruses can establish latency in transformed cells. The papillomaviruses are apparently more complex in their genome organization and encode many more regulatory proteins. Consequently, the papillomaviruses may be more highly evolved in this respect, but the extent to which the basic principles for attenuating replication in the different systems have diverged remains to be explored.

In earlier studies, frameshift mutations engineered into the M domain of the E1 ORF replicated at wild-type levels (3, 35), and we interpreted this to mean that this region of the ORF was not part of the replication protein (R). However, the data could not eliminate the possibility that reinitiation of translation from methionines encoded downstream of nt 849 occurred. Reinitiation of translation is more likely to provide a burst of protein synthesis in a transient replication assay in which a large amount of DNA is introduced into a given cell than in a stable assay in which the initial high dose of DNA is not maintained. Thus, mutants such as the E1Sma mutant which are significantly defective in the ability to produce either the 68-kDa or the M protein may allow for transient expression of a fragment of the 68-kDa protein. This fragment would thus potentially carry the enzymatic functions of the R gene product. It is clearly possible that the 5' domain of the E1 ORF encodes a regulatory domain which can be expressed as a 23-kDa protein and can also be part of a larger 68-kDa protein. This domain could be utilized to negatively regulate BPV-1 replication (see the report of Berg et al. [3]). Moreover, the enzymatic replication functions could be encoded by sequences 3' to the splice donor at nt 1235. Our discovery of a full-length E1 protein now raises two obvious questions. (i) Is the full-length E1 68-kDa protein the actual R protein defined by the mutations in the 3' end of the E1 ORF? (ii) Does another undetected R protein exist in virus-transformed cells? If the full-length 68-kDa protein is the

only E1-encoded replication protein, it then becomes necessary to attribute the reported complementation seen between the 3' and 5' mutants to an intracistronic complementation between carboxy-terminal fragments and the M protein. However, we should also point out that it was difficult in those complementation experiments to rule out physically undetectable but genetically significant recombination between the two mutants. These problems raised by our current comparisons of the genetic and biochemical studies need further clarification. Clearly the data presented here do not rule out a requirement for other undetected E1-encoded proteins whose genetic structures overlap those for the gene encoding the 68-kDa protein. Moreover, though our data support the notion that the methionine codon at nt 849 is the start codon for the 68-kDa protein, we cannot eliminate the possibility that its gene includes a small 3' exon. One approach to a resolution of this problem will be to ask how many BPV-1-encoded genes are both necessary and sufficient for plasmid replication.

ACKNOWLEDGMENTS

The first two authors contributed equally to the preparation of the manuscript.

We thank Ian Mohr for critical suggestions and help in the course of this work and Gary Bream for help with writing the manuscript.

This research was supported by Public Health Service grant CA42414 from the National Institutes of Health. M.L. is a postdoctoral fellow of the American Cancer Society and was supported by fellowship PF03150.

LITERATURE CITED

- Androphy, E. J., J. T. Schiller, and D. R. Lowy. 1985. Identification of the protein encoded by the E6 transforming gene of bovine papillomavirus. *Science* **230**:442-445.
- Baker, C. C. 1987. Sequence analysis of papillomavirus genomes, p. 321-385. In P. M. Howley (ed.), *The Papovaviridae*, vol. 2. The papilloma virus. Plenum Publishing Corp., New York.
- Berg, L., M. Lusky, A. Stenlund, and M. R. Botchan. 1986. Repression of bovine papilloma virus replication is mediated by a virally encoded *trans*-acting factor. *Cell* **46**:753-762.
- Bradley, M. K., T. F. Smith, R. H. Lathrop, D. M. Livingston, and T. A. Webster. 1987. Consensus topography in the ATP binding site of the simian virus 40 and polyomavirus large tumor antigens. *Proc. Natl. Acad. Sci. USA* **84**:4026-4030.
- Broker, T. R., and M. Botchan. 1986. Papillomaviruses: retrospective and prospective. *Cancer Cells* **4**:17-36.
- Buchkovich, K., L. A. Duffy, and E. Harlow. 1989. The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. *Cell* **58**:1097-1105.
- Burnett, S., U. Kiessling, and U. Pettersson. 1989. Loss of bovine papillomavirus DNA replication control in growth-arrested transformed cells. *J. Virol.* **63**:2215-2225.
- Choe, J., P. Vaillancourt, A. Stenlund, and M. Botchan. 1989. Bovine papillomavirus type 1 encodes two forms of a transcriptional repressor: structural and functional analysis of new viral cDNAs. *J. Virol.* **63**:1743-1755.
- Clertant, P., and F. Cuzin. 1982. Covalent affinity labeling by periodate-oxidized [α - 32 P]ATP of the large-T proteins of polyoma and SV40 viruses. *J. Biol. Chem.* **257**:6300-6305.
- Clertant, P., and I. Seif. 1984. A common function for polyoma virus large-T and papillomavirus E1 proteins? *Nature (London)* **311**:276-279.
- Colledge, W. H., W. D. Richardson, M. D. Edge, and A. E. Smith. 1986. Extensive mutagenesis of the nuclear location signal of simian virus 40 large-T antigen. *Mol. Cell Biol.* **6**:4136-4139.
- Cosman, D. J., and M. J. Tevethia. 1981. Characterization of a temperature-sensitive, DNA-positive nontransforming mutant of simian virus 40. *Virology* **112**:605-624.
- Dean, F. B., P. Bullock, Y. Murakami, C. R. Wobbe, L. Weissbach, and J. Hurwitz. 1987. Simian virus 40 (SV40) DNA replication: SV40 large T antigen unwinds DNA containing the SV40 origin of replication. *Proc. Natl. Acad. Sci. USA* **84**:16-20.
- Dvoretzky, I., R. Shober, S. K. Chattopadhyay, and D. Lowy. 1980. A quantitative *in vitro* focus assay for bovine papilloma virus. *Virology* **103**:369-375.
- Gluzman, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**:175-182.
- Grässer, F. A., K. Mann, and G. Walter. 1987. Removal of serine phosphates from simian virus 40 large T antigen increases its ability to stimulate DNA replication *in vitro* but has no effect on ATPase and DNA binding. *J. Virol.* **61**:3373-3380.
- Harlow, E., and D. Lane. 1988. *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith. 1984. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature (London)* **311**:33-38.
- Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith. 1984. A short amino acid sequence able to specify nuclear location. *Cell* **39**:499-509.
- Kaufman, R. J., M. V. Davies, V. K. Pathak, and J. W. B. Hershey. 1989. The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. *Mol. Cell. Biol.* **9**:946-958.
- Kawai, S., and M. Nishizawa. 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. *Mol. Cell. Biol.* **4**:1172-1174.
- Klausing, K., K.-H. Scheidtman, E. A. Baumann, and R. Knippers. 1988. Effects of *in vitro* dephosphorylation on DNA-binding and DNA helicase activities of simian virus 40 large tumor antigen. *J. Virol.* **62**:1258-1265.
- Korman, A. J., J. D. Frantz, J. L. Strominger, and R. C. Mulligan. 1987. Expression of human class II major histocompatibility complex antigens using retrovirus vectors. *Proc. Natl. Acad. Sci. USA* **84**:2150-2154.
- Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* **108**:229-241.
- Kramer, W., and H. Fritz. 1987. Oligonucleotide-directed construction of mutations via gapped duplex DNA. *Methods Enzymol.* **154**:350-367.
- Lambert, P. F., C. C. Baker, and P. M. Howley. 1988. The genetics of bovine papillomavirus type 1. *Annu. Rev. Genet.* **22**:235-258.
- Lambert, P. F., and P. M. Howley. 1988. Bovine papillomavirus type 1 E1 replication-defective mutants are altered in their transcriptional regulation. *J. Virol.* **62**:4009-4015.
- Lanford, R. E., and J. S. Butel. 1984. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37**:801-813.
- Lanford, R. E., P. Kanda, and R. C. Kennedy. 1986. Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. *Cell* **46**:575-582.
- Lautenberger, J. A., A. Seth, C. Jorczyk, and T. S. Papas. 1984. Useful modifications of the *Escherichia coli* expression plasmid pJL6. *Gene Anal. Tech.* **1**:63-66.
- Law, M.-F., D. R. Lowy, I. Dvoretzky, and P. M. Howley. 1981. Mouse cells transformed by bovine papillomavirus contain only extrachromosomal viral DNA sequences. *Proc. Natl. Acad. Sci. USA* **78**:2727-2731.
- Lusky, M., and M. R. Botchan. 1981. Inhibition of SV-40 replication in simian cells by specific pBR322 DNA sequences. *Nature (London)* **293**:79-81.
- Lusky, M., and M. R. Botchan. 1984. Characterization of the bovine papilloma virus plasmid maintenance sequences. *Cell* **36**:391-401.
- Lusky, M., and M. R. Botchan. 1985. Genetic analysis of bovine papillomavirus type 1 *trans*-acting replication factors. *J. Virol.* **53**:955-965.
- Lusky, M., and M. R. Botchan. 1986. A bovine papillomavirus type 1-encoded modulator function is dispensable for transient viral replication but is required for establishment of the stable

- plasmid state. *J. Virol.* **60**:729–742.
36. McVey, D., L. Brizuela, I. Mohr, D. R. Marshak, Y. Gluzman, and D. Beach. 1989. Phosphorylation of large tumour antigen by cdc2 stimulates SV40 DNA replication. *Nature (London)* **341**: 503–507.
 37. Mohr, I. J., Y. Gluzman, M. P. Fairman, M. Strauss, D. McVey, B. Stillman, and R. D. Gerard. 1989. Production of simian virus 40 large tumor antigen in bacteria: altered DNA-binding specificity and DNA-replication activity of underphosphorylated large tumor antigen. *Proc. Natl. Acad. Sci. USA* **86**:6479–6483.
 38. Mohr, I. J., B. Stillman, and Y. Gluzman. 1987. Regulation of SV40 DNA replication by phosphorylation of T antigen. *EMBO J.* **6**:153–160.
 39. Nurse, P. 1990. Universal control mechanism regulating onset of M-phase. *Nature (London)* **344**:503–508.
 40. Peabody, D. S. 1989. Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.* **264**:5031–5035.
 41. Platt, E. J., K. Karlsen, A. Lopez-Valdivieso, P. W. Cook, and G. L. Firestone. 1986. Highly sensitive immunoadsorption procedure for detection of low-abundance proteins. *Anal. Biochem.* **156**:126–135.
 42. Richardson, W. D., B. L. Roberts, and A. E. Smith. 1986. Nuclear location signals in polyoma virus large-T. *Cell* **44**:77–85.
 43. Rosenberg, A. H., B. N. Lade, D.-S. Chui, S.-W. Lin, J. J. Dunn, and F. W. Studier. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**:125–135.
 44. Rossmann, M. G., A. Lilijas, C. I. Braden, and L. J. Banaszak. 1975. Evolutionary and structural relationships among dehydrogenases, p. 61–102. *In* P. D. Boyer (ed.), *The enzymes*, vol. 11. Academic Press, Inc., New York.
 45. Rossmann, M. G., D. Moras, and K. W. Olsen. 1974. Chemical and biological evolution of a nucleotide-binding protein. *Nature (London)* **250**:194–199.
 46. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 47. Schiller, J. T., E. Kleiner, E. J. Androphy, D. R. Lowy, and H. Pfister. 1989. Identification of bovine papillomavirus E1 mutants with increased transforming and transcriptional activity. *J. Virol.* **63**:1775–1782.
 48. Seedorf, K., T. Oltersdorf, G. Krämmer, and W. Röwekamp. 1987. Identification of early proteins of the human papilloma viruses type 16 (HPV 16) and type 18 (HPV 18) in cervical carcinoma cells. *EMBO J.* **6**:139–144.
 49. Seif, I., G. Khoury, and R. Dhar. 1980. A rapid enzymatic DNA sequencing technique: determination of sequence alterations in early simian virus 40 temperature sensitive and deletion mutants. *Nucleic Acids Res.* **8**:2225–2240.
 50. Shaw, S. B., and P. Tegtmeyer. 1981. Binding of dephosphorylated A protein to SV40 DNA. *Virology* **115**:88–96.
 51. Sigel, M. B., Y. N. Sinha, and W. P. VanderLaan. 1983. Production of antibodies by inoculation into lymph nodes. *Methods Enzymol.* **93**:3–12.
 52. Simmons, D. T. 1986. DNA-binding region of the simian virus 40 tumor antigen. *J. Virol.* **57**:776–785.
 53. Smith, A. E., D. Kalderon, B. L. Roberts, W. D. Colledge, M. Edge, P. Gillett, A. F. Markham, E. Paucha, and W. D. Richardson. 1985. The nuclear location signal. *Proc. R. Soc. London B* **226**:43–58.
 54. Stahl, H., P. Dröge, and R. Knippers. 1986. DNA helicase activity of SV40 large tumor antigen. *EMBO J.* **5**:1939–1944.
 55. Stenlund, A., and M. Botchan. 1990. The E2 *trans*-activator can act as a repressor by interfacing with a cellular transcription factor. *Genes Dev.* **4**:123–136.
 56. Thomas, C. A., K. J. Berns, Jr., and T. Kelly, Jr. 1966. Isolation of high molecular weight DNA from bacteria and cell nuclei, p. 535. *In* G. L. Cantoni and D. R. Davies (ed.), *Procedures in nucleic acid research*. Harper & Row, Publishers, Inc., New York.
 57. Thorner, L., N. Bucay, J. Choe, and M. Botchan. 1988. The product of the bovine papillomavirus type 1 modulator gene (*M*) is a phosphoprotein. *J. Virol.* **62**:2474–2482.
 58. Wigler, M., A. Pellicer, S. Silverstein, and R. Axel. 1978. Biochemical transfer of single copy eukaryotic genes using total cellular DNA as a donor. *Cell* **14**:725–731.
 59. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.