JOURNAL OF VIROLOGY, Feb. 1996, p. 1117–1124 Vol. 70, No. 2
0022-538X/96/$04.00+0

The Bovine Papillomavirus Type 1 E2 Transactivator and Repressor Proteins Use Different Nuclear Localization Signals

MARIO H. SKIADOPOULOS AND ALISON A. MCBRIDE*

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 20 September 1995/Accepted 30 October 1995

The E2 gene of bovine papillomavirus type 1 encodes at least three nuclear phosphoproteins that regulate viral transcription and DNA replication. All three proteins have a common C-terminal domain that has DNA-binding and dimerization activities. A basic region in this domain forms an alpha helix which makes direct contact with the DNA target. In this study, it is shown that in addition to its role in DNA binding, this basic region functions as a nuclear localization signal both in the E2 DNA-binding domain and in a heterologous protein. Deletion of this signal sequence resulted in increased accumulation of the E2 transactivator and repressor proteins in the cytoplasm, but nuclear localization was not eliminated. In the full-length transactivator protein, another signal, present in the N-terminal transactivation domain, is used for transport to the nucleus, and the C-terminal nuclear localization signal(s) are masked. The use of different nuclear localization signals could potentially allow differential regulation of the subcellular localization of the E2 transactivator and repressor proteins at some stage in the viral life cycle.

The E2 gene of bovine papillomavirus type 1 encodes several proteins that are required for regulation of transcription of early viral promoters and for replication of the viral genome (reviewed in references 12, 15, and 27). A 48-kDa polypeptide (E2-TA) is a transcriptional transactivator and contains an amino-terminal transactivation domain and a carboxy-terminal DNA-binding and dimerization domain, separated by a hinge region. Two smaller E2 polypeptides, E2-TR (31 kDa) and E8/E2 (28 kDa), are transcribed from different promoters and have common C-terminal sequences with E2-TA. The initiating methionine of E2-TR is located at codon 162 of the E2 open reading frame, and E8/E2 contains 11 amino acids from the E8 open reading frame fused to codon 206 of E2 by use of alternative promoter and splicing sites. All three protein species can form stable homodimers and heterodimers, which bind to specific E2 DNA-binding sites ($ACCN_6GGT$). The smaller E2 proteins can antagonize the function of the E2 transactivator and act as transcriptional repressors (3, 7, 18). Two mechanisms for this repression have been postulated. In one model, the E2 repressors form inactive heterodimers with E2-TA: these heterodimers could bind to E2 DNA-binding sites within the BPV-1 enhancer elements but would be unable to activate the promoters. In the other model, the E2-TR and E8/E2 proteins bind to E2 DNA-binding sites as homodimers and block binding of E2-TA. The E2 transactivator is also required for viral DNA replication (42). E2-TA forms a stable complex with the E1 protein, and these proteins bind cooperatively to the origin of replication, which contains an E1-binding site flanked by E2-binding sites (33, 39, 41, 43, 45, 46). The E2-TA protein also enhances DNA replication by alleviating nucleosomal repression in the replication origin (20) and by interacting with replication proteins such as replication protein A (19). The ratio of the E2 transactivator and repressor proteins varies throughout the cell cycle, and increased amounts of the E2 transactivator are present during S phase (47).

* Corresponding author. Mailing address: Rm. 137, Bldg. 4, National Institutes of Health, 4 Center Dr. MSC 0455, Bethesda, MD 20892-0455. Phone: (301) 496-1370. Fax: (301) 480-1560. Electronic mail address: alison_mcbride@nih.gov.

Cell fractionation and indirect immunofluorescence studies have shown that the E2 proteins are localized predominantly in the nucleus (4, 16). Proteins that function in the nucleus must first be transported from the cytoplasm, through the nuclear membrane, and into the nucleus. Several nuclear transport models have been proposed (reviewed in references 10 and 37). Small proteins are able to enter the nucleus by diffusion if their size is within the exclusion limit of the nuclear pores, which has been estimated to be between 40 and 60 kDa. However, proteins smaller than this may still be transported by an active-transport mechanism. Active transport requires a nuclear localization signal (NLS) and is mediated by an NLS-binding protein that, along with other cellular factors, shuttles proteins across the nuclear membrane in an energy-dependent manner. Proteins as small as the H2B histone protein (15 kDa) have been shown to enter the nucleus by NLS-mediated active transport (30). Proteins that do not have an NLS can enter the nucleus by "piggyback" transport, which requires interaction with a protein that does contain an NLS.

An NLS consensus sequence has not yet been established, other than a preponderance of basic residues within a relatively short stretch of amino acids (9). Two major classes of NLS have been described: the simian virus 40 (SV40) large T-antigen-type NLS (PKKKRKV), which consists of several basic amino acids in close proximity, and the bipartite NLS, which consist of two basic regions separated by a spacer region of variable length (usually 10 to 12 amino acids [aa]) (32). Proteins may have more than one NLS, and one or all may be required for nuclear localization.

The ability to control entry of a protein into the nucleus appears to be a regulatory mechanism for many proteins. Phosphorylation at sites adjacent to NLSs can regulate nuclear transport. For example, phosphorylation at casein kinase II sites near the SV40 T-antigen NLS affects the rate of entry of T antigen into the nucleus (17), and phosphorylation of the yeast SWI5 protein regulates transport of this protein into the nucleus in a cell cycle-dependent manner (29). Nuclear transport can also be regulated by masking the NLS sequence. This can be achieved by interaction with another polypeptide, such as occurs with the NF-κB protein (2), or by an intramolecular

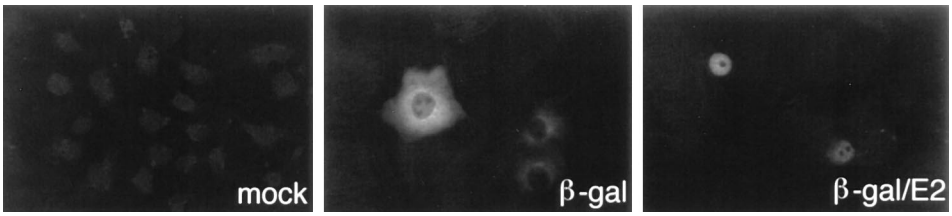

FIG. 1. A β-galactosidase–E2 fusion protein is transported to the nucleus. β-Galactosidase-expressing plasmids were transfected into Cos-7 cells, and subcellular localization was determined by indirect immunofluorescence with antiserum against β-galactosidase. mock, no DNA; βgal, pCH110; βgal/E2, pCH110-E2.

conformational change, such as occurs in the human heat shock factor 2 (36).

In this study, we have used indirect immunofluorescence to identify the NLSs of the E2 proteins. We have determined that the E2-TA protein utilizes an NLS located in the unique N-terminal transactivation domain. The E2-TR protein does not contain this sequence and uses other signals to enter the nucleus. However, although these C-terminal NLS sequences are also present in the E2 transactivator, they appear to be masked and do not function in nuclear transport.

MATERIALS AND METHODS

Cell culture and viruses. Cos-7 and CMT4 cells (11) were grown in Dulbecco's minimal essential medium supplemented with 10% fetal calf serum. Recombinant SV40 PAVA E2 virus was produced in CMT4 cells, as described previously (34).

Plasmids and mutagenesis. The recombinant PAVA virus expressing the E2 protein, pPAVA$_{kz}$E2, has been described previously (pSB-E2$_{kz}$ [23]). pPAVA E2-TR consists of pPAVA with an amber mutation (TAC to TAG) at nucleotide 3082 (34). Plasmids C59-E2$_{kz}$, C59-E2$_{dl1-52}$, C59-E2$_{dl92-161}$, C59-E2$_{dl158-282}$, C59-E2$_{dl220-309}$, C59-E2$_{dl361-410}$ (E2$_{1-360}$) and C59-E2$_{dl211-410}$ (E2$_{1-210}$) have been described previously (24). *Bst*EII (nucleotide 2405) to *Bst*XI (nucleotide 3889) fragments containing the E2 sequences were cloned from the C59 plasmids into pPAVA-E2 to generate pPAVA-E2$_{dl1-52}$, pPAVA-E2$_{dl92-161}$, pPAVA-E2$_{dl158-282}$, pPAVA-E2$_{dl220-309}$, pPAVA-E2$_{dl361-410}$, and pPAVA-E2$_{dl211-410}$. pPAVA-E2$_{290-410}$ was generated by replacing the *Bst*EII-*Bst*XI fragment of pPAVA$_{kz}$E2 with that from pTZmodE2$_{290-410}$. pTZmodE2$_{290-410}$ was generated from pTZE2$_{290-410}$ (24) by cleaving with *Sty*I (nucleotide 3535) and *Pfl*MI (nucleotide 3683) and replacing the released fragment with five pairs of overlapping oligonucleotides. Two of the oligonucleotides generated new unique restriction sites in the E2 gene without changing the E2 amino acid sequence (*Dra*II, A to C at nucleotide 3585; *Cla*I, C to A at nucleotide 3633). To generate deletion (*dl*BR3) or substitution (*sb*BR3) mutations, double-stranded oligonucleotides encoding aa 325 to 359 were ligated between the new *Dra*II and *Pfl*MI (nucleotide 3683) sites of pTZmodE2$_{290-410}$. In pTZE2$_{290-410/dl\text{BR3}}$, the oligonucleotides generated a deletion of aa 339 to 352, and in pTZE2$_{290-410/sb\text{BR3}}$, residues K-339, R-342, R-344, K-346, K-347, R-350, and R-352 were substituted to asparagine by replacing each arginine and lysine codon with the nucleotides AAC. The E2 DNA-binding domain containing these mutations was subcloned into pPAVA$_{290-410}$, pPAVA$_{kz}$E2, and C59-E2-TR.

C59-E2-TR (aa 162 to 410), a plasmid that expresses the E2-TR cDNA, has been described previously (p1153 [18]). C59-E2-TR$_{dl284-410}$ was generated by inserting a translation termination linker (5′ TTAGTTAACTAA 3′) into the *Asp*718 (nucleotide 3460) site of C59-E2-TR.

Plasmids pCMVT7-E2$_{dl41-50}$, pCMVT7-E2$_{dl101-110}$, and pCMVT7-E2$_{dl111-120}$ were provided by J.-M. Gauthier, Pasteur Institute, Paris, France. These constructs have the 10 aa indicated deleted and replaced with a linker encoding a *Sac*I site and amino acids GlyAlaLeu. The *Bst*EII-*Bst*XI E2 insert of these plasmids was cloned into pPAVA$_{kz}$E2 to generate pPAVA-E2$_{dl41-50}$, pPAVA-E2$_{dl101-110}$, and pPAVA-E2$_{dl111-120}$, respectively. pPAVA-E2$_{dl101-120}$ and pPAVA-E2$_{dl41-120}$ were created by subcloning the *Sac*I-*Bst*XI fragment of pPAVA-E2$_{dl111-120}$ into pPAVA-E2$_{dl101-110}$ and pPAVA-E2$_{dl41-50}$, respectively.

pPAVA$_{kz}$E2-TA, pPAVA$_{kz}$E2TA$_{dl101-110}$, pPAVA$_{kz}$E2TA$_{dl111-120}$, pPAVA$_{kz}$E2TA$_{dl101-120}$, pPAVA$_{kz}$E2TA$_{dl41-120}$, pPAVA$_{kz}$E2TA$_{dl\text{BR3}}$, and pPAVA$_{kz}$E2TA$_{sb\text{BR3}}$, were designed to express the E2-TA protein only. The initiating methionine of E2-TR in these constructs was changed to an isoleucine by mutating nucleotide 3093 from ATG to ATC with the Transformer mutagenesis kit (Clontech). All mutations were verified by DNA sequencing.

β-Galactosidase fusion protein expression vectors were created by inserting double-stranded oligonucleotides, specifying the sequence for BR1 (ARKKG), BR2 (EPKRCFKKGARV), BR3 (NQVKCYRFRVKKNHRHRY), or the T-antigen NLS (PKRKKKV) into the unique *Asp*718 site of pCH110 (Pharmacia) such that the inserted sequences were fused in frame near the N-terminal end of β-galactosidase. To create the β-galactosidase–E2 fusion expression vector, a PCR-generated E2 fragment was ligated into the *Eco*RI site (nucleotide 3500) of pCH110, such that the E2 coding sequence (codons 3 to 410) was fused in frame near the C-terminal end of β-galactosidase.

Transient expression and immunofluorescence. Cos-7 cells were plated onto glass slides 16 h before infection or transfection. For PAVA virus E2 expression, Cos-7 cells were infected with virus at a high multiplicity of infection and analyzed for E2 expression after 40 to 44 h. For transfections, Cos-7 cells grown on slides in 100-mm-diameter plates were transfected with 20 μg of plasmid DNA by using DEAE-dextran or Lipofectamine (Gibco/BRL). Transfection experiments were repeated several times by both methods. Transfected cells were analyzed approximately 40 h posttransfection. Cells were fixed for 30 min in 3.7% formaldehyde solution in phosphate-buffered saline (PBS) and permeabilized with 0.1% Triton X-100 in PBS for 4 min. Slides were incubated in PBS block solution (0.25% gelatin plus 0.25% bovine serum albumin [BSA] in PBS) for 1 h at room temperature. Mouse monoclonal anti-E2 antibodies B201 and B202 (provided by Elliot Androphy) were added at dilutions of 1:10 and 1:100 in PBS, respectively. Affinity-purified rabbit polyclonal antiserum SRQE was used at a 1:30 dilution in PBS (28). Antibodies against β-galactosidase were used at a 1:100 dilution (Sigma Immunochemicals). Slides were incubated for 1 h with the primary antibody, washed three times with PBS, and incubated for 1 h with goat anti-mouse immunoglobulin or donkey anti-rabbit immunoglobulin G conjugated to fluorescein isothiocyanate (1:100 dilution; Jackson Immunochemicals). Following three washes in PBS, slides were mounted in Vectashield mounting fluid (Vector Laboratories). Immunofluorescence was detected and photographed with a Bio-Rad MRC600 confocal laser scanning imaging system.

RESULTS

A β-galactosidase–E2 fusion protein is transported to the nucleus. Theoretically, monomeric E2 proteins are small enough to enter the nucleus by diffusion. To determine whether the E2 transactivator is transported to the nucleus by active transport, as opposed to diffusion, a plasmid that expresses a β-galactosidase–E2-TA fusion protein was generated. Cos-7 cells were transfected with this plasmid, and the subcellular location of the protein was examined by using indirect immunofluorescence and anti-β-galactosidase antibodies. As shown in Fig. 1, fusion of the E2 open reading frame to the C-terminal end of β-galactosidase causes this normally cytoplasmic protein to enter the nucleus. Therefore, the E2-TA protein contains sequences that mediate active transport across the nuclear membrane.

DNA binding and dimerization are not required for localization. The E2 proteins form dimers through the C-terminal DNA-binding domain. To analyze whether DNA binding and polypeptide dimerization are required for nuclear localization or for nuclear retention, the subcellular localization of E2 proteins with point mutations in the DNA-binding domain was examined. These proteins have amino acid substitutions in the DNA-binding domain and are able to bind DNA and dimerize (K339R, C340G), are defective only in DNA binding (R344K), or are defective in both properties (V345R) (5, 26). Cos-7 cells were infected with PAVA viruses that express these E2 proteins, and their subcellular location was determined by indirect immunofluorescence. As shown in Fig. 2, the wild-type E2 protein shows strong nuclear immunofluorescence that is excluded from the nucleolus. All of the mutated E2 proteins are also located in the nucleus, showing that the DNA-binding and

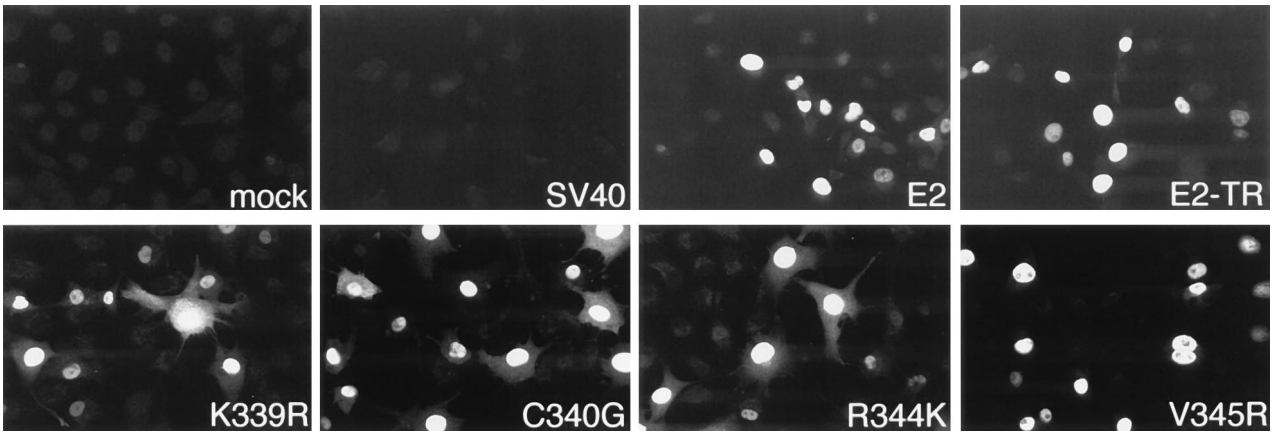

FIG. 2. DNA binding and dimerization are not required for localization. Cos-7 cells were infected with recombinant viruses expressing the indicated mutated E2 proteins. Subcellular localization was determined by indirect immunofluorescence with B201 E2-specific antiserum. mock, no virus; SV40, SV40; E2, $PAVA_{kz}E2$; E2-TR, $PAVA_{kz}E2$-TR; K339R, $PAVA_{kz}E2_{K339R}$; C340G, $PAVA_{kz}E2_{C340G}$; R344K, $PAVA_{kz}E2_{R344K}$; V345R, $PAVA_{kz}E2_{V345R}$.

dimerization properties are not required for either localization or retention of the E2 protein in the nucleus.

The basic region in the DNA-binding domain of BPV-1 E2 can function as an NLS. The primary amino acid sequence of the E2 protein was examined for potential NLSs, and three regions that contain basic residues in close proximity were identified (Fig. 3A). These regions were designated BR1, BR2, and BR3. Only the third region (BR3), which is present in the

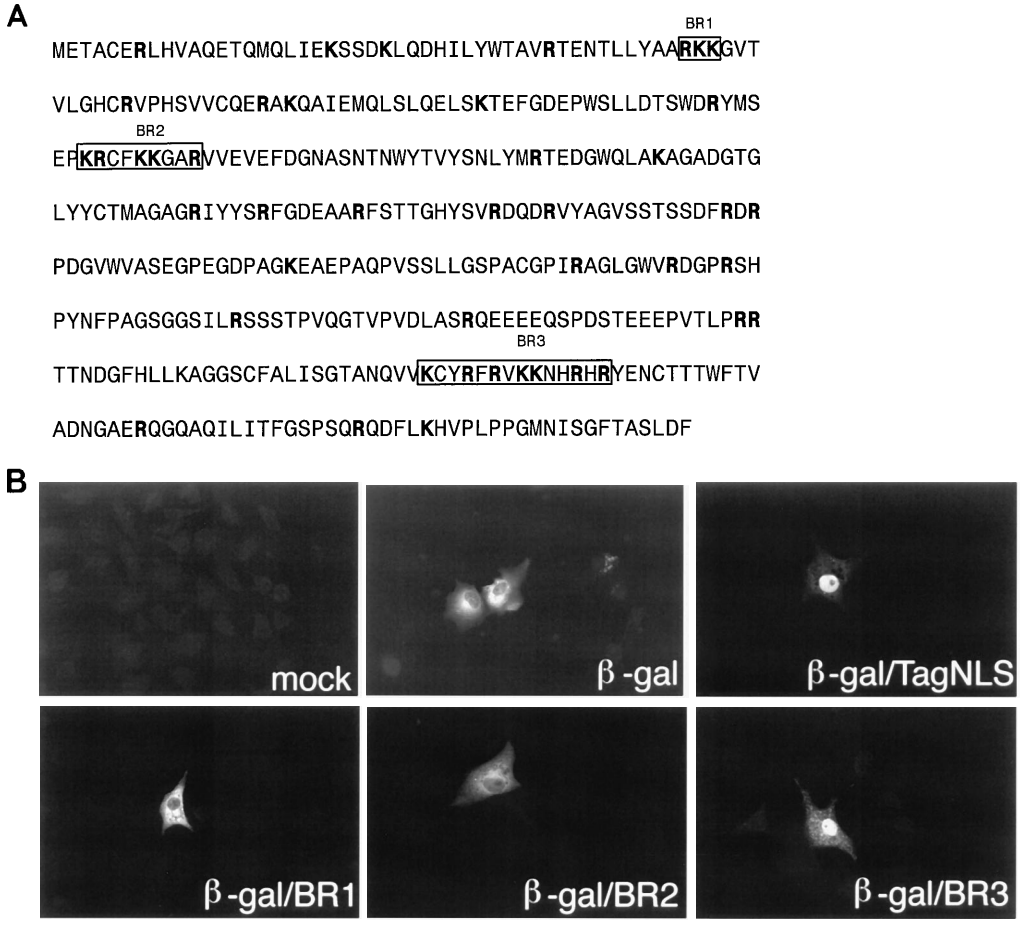

FIG. 3. (A) Location of basic regions in the E2 proteins. Amino acid sequence of the BPV-1 E2-TA protein with the basic regions BR1, BR2, and BR3 boxed. Basic residues are shown in boldface type. (B) The basic region in the DNA-binding domain of BPV-1 E2 can function as an NLS. β-Galactosidase-expressing plasmids were transfected into Cos-7 cells, and subcellular localization was determined by indirect immunofluorescence with antiserum against β-galactosidase. mock, no DNA; β-gal, pCH110; β-gal/TagNLS, pCH110-TagNLS β-gal/BR1, pCH110-BR1; β-gal/BR2, pCH110-BR2; β-gal/BR3, pCH110-BR3.

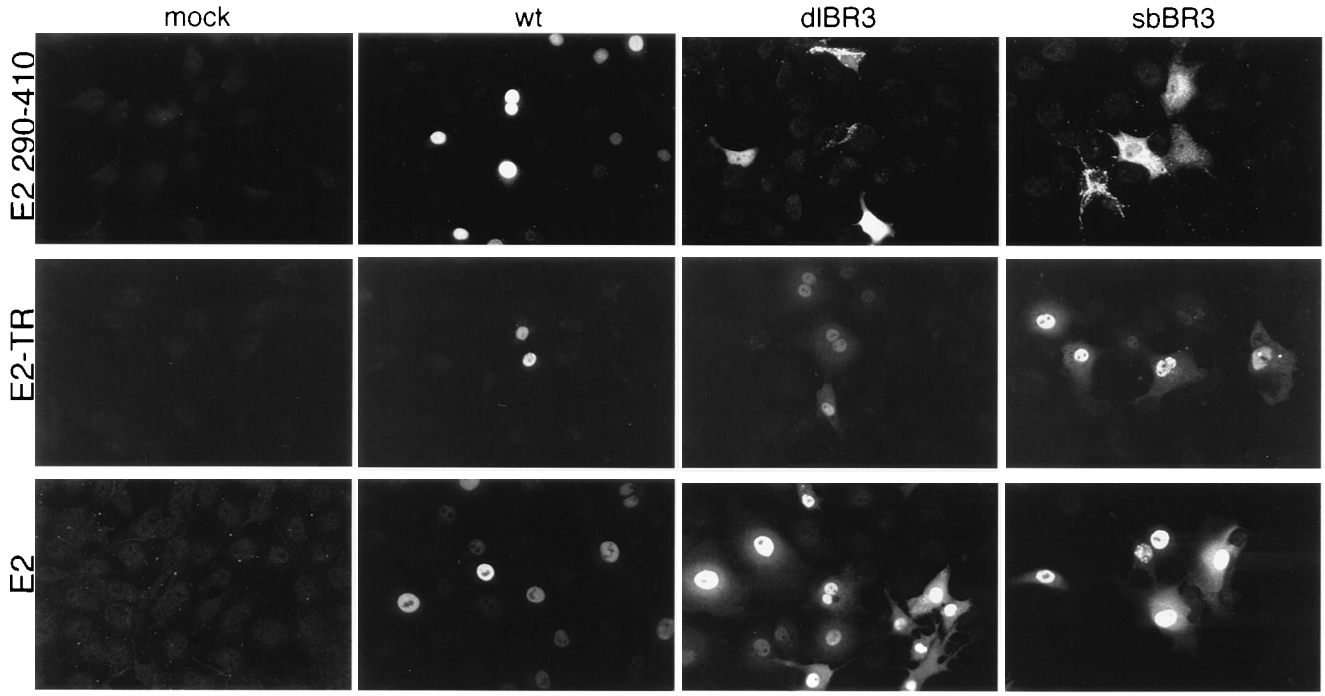

FIG. 4. Effect of deletion or substitution of BR3 on E2 proteins. Cos-7 cells were infected (E2-TA and $E2_{290–410}$) or transfected (E2-TR) with recombinant viruses or DNAs expressing $E2_{290–410}$, E2-TR, or E2-TA proteins with deletion or substitution of BR3, as indicated. E2 localization was determined by indirect immunofluorescence with the E2-specific antisera SRQE($E2_{290–410}$) or B201 (E2-TA and E2-TR). wt, wild type.

DNA-binding domain, is common to all three species of the BPV-1 E2 polypeptides. This region has been shown to form an alpha helix that directly contacts the E2 DNA-binding site (13) and is highly conserved among all papillomavirus E2 proteins.

To determine whether the basic-residue-rich regions of the E2 protein could function as NLSs, sequences encoding them were fused in frame near the N-terminal end of β-galactosidase. The SV40 large T-antigen NLS (PKKKRKV) was used as a positive control. Cos-7 cells were transfected with the β-galactosidase fusion vectors, and the subcellular location was determined by using antibodies against the β-galactosidase protein. As shown in Fig. 3B, only the BR3 sequence from the DNA-binding domain of the E2 protein and the large T-antigen NLS were able to direct β-galactosidase to the nucleus. Therefore, the basic region of the BPV-1 E2 DNA-binding domain is an NLS that can mediate transport of a heterologous protein into the nucleus.

BR3 is required for nuclear localization of an E2 DNA-binding domain polypeptide but is not absolutely required for nuclear localization of the E2-TR and E2-TA proteins. To examine whether BR3 can function as an NLS within the context of the E2 protein, this sequence was mutated in a polypeptide that contains the DNA-binding domain of E2 ($E2_{290–410}$ [see Fig. 5A]). This 122-aa polypeptide, which is common to all three species of E2, can form dimers and bind DNA (24, 28) and can repress transcription in vivo (22). As shown in Fig. 4, this polypeptide is localized to the nucleus. When BR3 is deleted (*dl*BR3) in this protein or when all the lysine and arginine residues in BR3 are converted to asparagine (*sb*BR3), localization to the nucleus is disrupted. Cells expressing the $E2_{290–410/sb\text{BR3}}$ or $E2_{290–410/dl\text{BR3}}$ proteins displayed either a punctate pattern in the cytoplasm, a weak nuclear staining, or staining that was distributed throughout the cytoplasm and the nucleus. $E2_{290–410/dl\text{BR3}}$ and $E2_{290–410/sb\text{BR3}}$ proteins were examined for their ability to bind DNA or form dimers and were found to be completely defective for these activities (data not shown). These results demonstrate that BR3 is required for nuclear localization within the context of the E2 C-terminal domain.

To establish whether BR3 was necessary and sufficient for nuclear transport of the E2-TA and E2-TR proteins, the effect of the BR3 deletion and substitution mutations in the background of these full-length proteins was analyzed. As shown in Fig. 4, nuclear localization was affected but not completely blocked by the substitution or deletion of BR3. Strong nuclear staining was detected in Cos-7 cells transfected or infected with vectors expressing these proteins, but a reproducible increase in cytoplasmic staining was also observed. Thus, the basic region in the DNA-binding domain is required for efficient nuclear localization of the E2 proteins, but other signals may also be involved.

Analysis of the subcellular localization of E2 proteins containing deletions. The ability of the E2-TA and E2-TR *dl*BR3 and *sb*BR3 proteins to localize to the nucleus suggested that the E2 proteins contained additional NLSs. To examine this possibility, the subcellular location of a series of E2 proteins with deletions in the transactivation domain, DNA-binding domain, or hinge region was examined by indirect immunofluorescence. As shown in Fig. 5, deletion of the extreme C-terminal region of E2 (aa 361 to 410) or a portion of the hinge region of E2 (aa 158 to 282; aa 220 to 309) had no apparent effect on localization of E2 to the nucleus, suggesting that these sequences are not required for nuclear transport. These deletions do not remove basic regions BR1, BR2, or BR3.

Truncation of the entire C-terminal half of the E2-TR (E2-$TR_{dl284–410}$) or E2-TA ($E2_{dl211–410}$) proteins altered nuclear localization of E2. These proteins localized mainly in the nu-

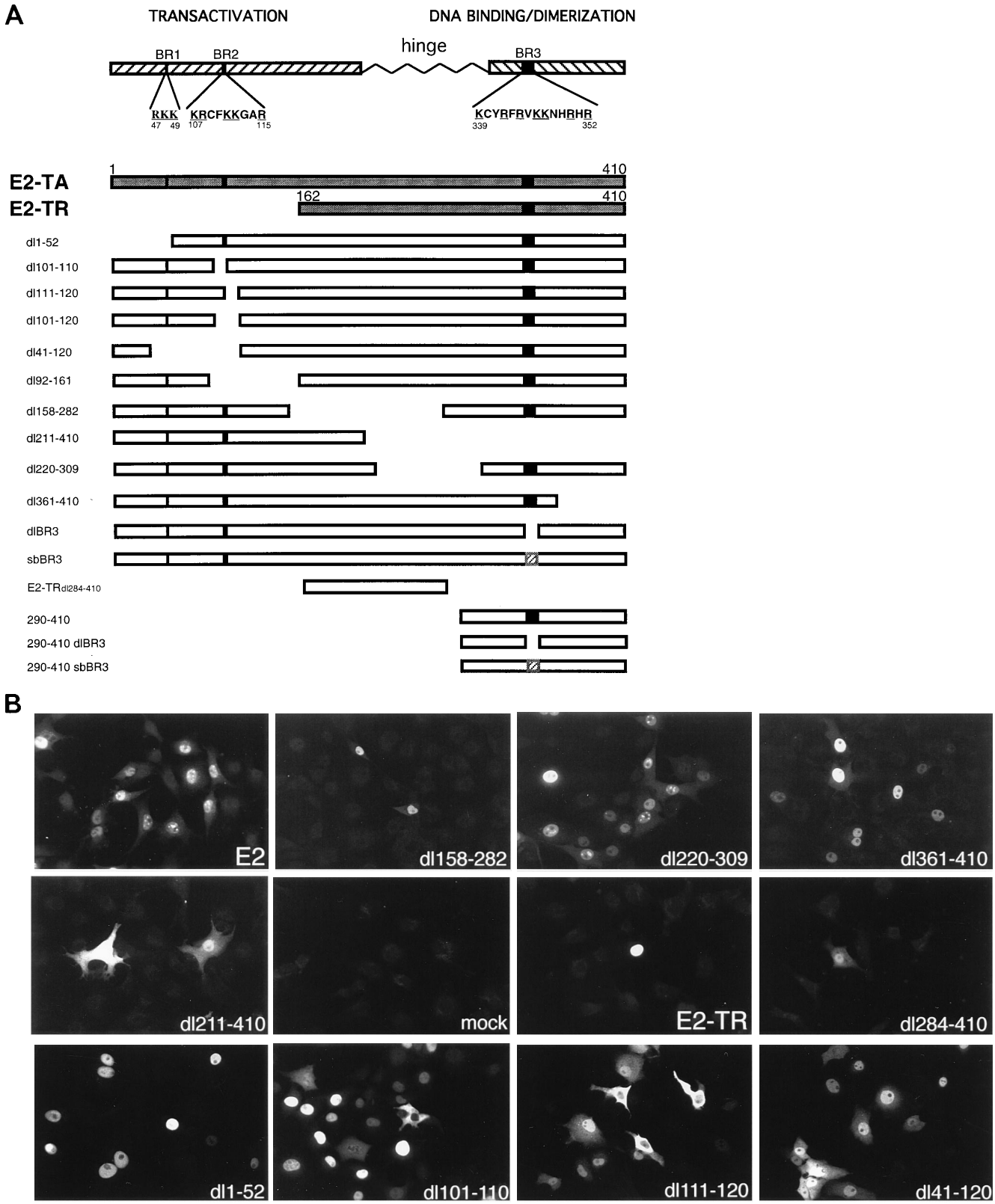

FIG. 5. (A) Deletion mutations in the E2 protein. A diagram of the E2-TA and E2-TR proteins is shown at the top, with the positions and sequences of the BR1, BR2, and BR3 regions indicated. The structures of the mutated E2 proteins are shown below. (B) Deletion analysis of the E2 proteins identifies additional sequences affecting nuclear localization. Cos-7 cells were infected with recombinant PAVA viruses expressing the mutated E2 proteins indicated. E2 localization was determined by indirect immunofluorescence with the E2-specific antisera B202 (E2 and *dl*158–282) or B201 (*dl*220–309, *dl*361–410, *dl*211–410, mock, E2-TR, *dl*283–410, *dl*1–52, *dl*101–110, *dl*111–120, and *dl*41–120).

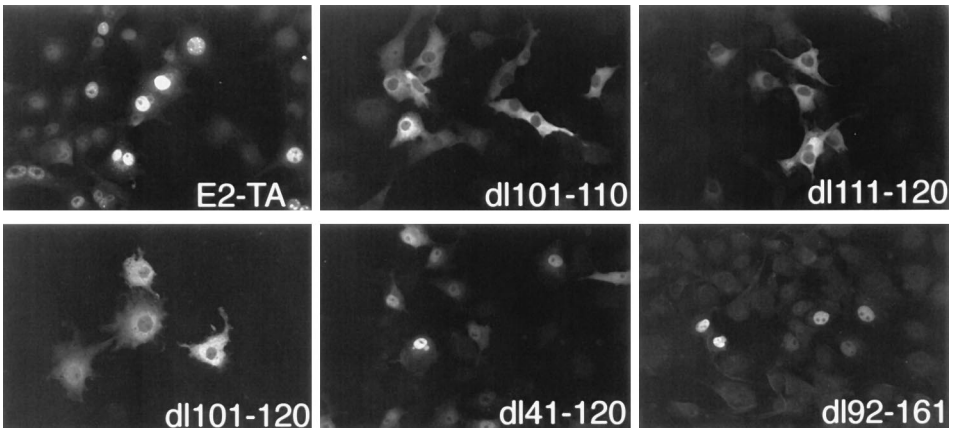

FIG. 6. E2-TA contains an NLS in the N-terminal domain, and C-terminal NLSs are masked. Cos-7 cells were infected with recombinant PAVA E2-TA viruses expressing the mutated E2 proteins indicated. The initiation codon of E2-TR was mutated in these viruses, so that only the E2-TA protein could be expressed. E2 localization was determined by indirect immunofluorescence with the E2-specific antiserum B201.

cleus but also accumulated in the cytoplasm, indicating that sequences required for complete nuclear localization had been perturbed. However, the molecular weight of these polypeptides would be predicted to be below the exclusion limits of the nuclear pore, and passive diffusion could occur.

Deletion of the extreme N-terminal end of E2-TA ($E2_{dl1-52}$), which includes BR1, also did not affect nuclear localization. However, deletion of BR2 ($E2_{dl101-120}$) or portions of BR2 ($E2_{dl111-120}$ or $E2_{dl101-110}$) resulted in a mixed population of cells with respect to the subcellular distribution of the E2 protein. The majority of cells displayed nuclear staining, but a significant proportion stained both the nucleus and cytoplasm or the cytoplasm exclusively. Since BR2 is present in the unique transactivation domain, deletion of this region probably affects the localization of only the E2-TA protein. Because the E2-TR protein is expressed from a promoter contained within the E2 open reading frame, the PAVA viruses can express both the E2-TA and E2-TR proteins. This could explain the mixed distribution of E2 staining observed in cells infected with these viruses. This analysis is further complicated by the ability of the E2 proteins to form heterodimers between E2-TA and E2-TR species.

The E2-TA protein contains an NLS in the N-terminal transactivation domain. To examine the role of BR2 in the E2-TA protein, the initiating methionine (codon ATG) of E2-TR was mutated to isoleucine (codon ATC), thereby eliminating the expression of E2-TR from the recombinant viruses. This mutation has no effect on the transactivation or replication properties of the E2 proteins (42). As shown in Fig. 6, E2-TA expressed in the absence of E2-TR is nuclear, and deletion of aa 101 to 110 and/or 111 to 120 resulted in accumulation of the protein in the cytoplasm. This demonstrates that the BR2 region may function as an NLS that is necessary for nuclear localization of the E2-TA protein. In addition, we found that E2-TA proteins with a proline-to-glycine substitution at residue 106 accumulate in the cytoplasm (38).

A nuclear targeting signal(s) is masked in the E2 transactivator. Small deletions in BR2 (*dl*101–110, *dl*111–120, or *dl*101–120) resulted in cytoplasmic accumulation of the E2-TA protein. However, E2-TA proteins containing these deletions still encode the BR3 NLS and any other NLSs utilized by the E2-TR repressor, yet these signals do not appear to be used. This finding suggests that in the background of the E2-TA protein, the C-terminal signals are normally masked. Furthermore, proteins with larger deletions in the transactivation domain (*dl*41–120 or *dl*92–161) were targeted to the nucleus in many of the cells examined (Fig. 6). This suggests that the larger deletions are disrupting protein conformation and are unmasking the C-terminal NLS. A further deletion of BR3 within the background of E2-TA *dl*41–120 perturbs nuclear localization, which may indicate that BR3 is, at least partially, responsible for nuclear uptake of the *dl*41–120 E2 protein (data not shown). These findings are summarized in the diagram in Fig. 7.

DISCUSSION

We have identified two regions in the BPV-1 E2 proteins that are required for efficient localization to the nucleus. The E2-TA polypeptide is completely defective for nuclear uptake when a basic region (BR2) in the transactivation domain is deleted. Deletion of either aa 101 to 110 or 111 to 120 resulted in cytoplasmic localization of E2-TA. This region contains five positively charged residues flanked by proline and valine (PKRCFKKGARV) and bears some similarity to the SV40 large T-antigen NLS (PKKKRKV). However, unlike the T-antigen NLS, this sequence was unable to direct β-galactosidase to the nucleus. This inability to direct a cytoplasmic protein to the nucleus has also been found with the NLSs of other proteins (8), and it has been suggested that this may be due to improper presentation of the NLS in the context of the heterologous protein. Alternatively, BR2 may only be part of a larger NLS in E2 and may be unable to function alone. Deletion of a larger region around BR2 in the E2-TA transactivation domain restored nuclear uptake of the E2 protein. This could indicate that the larger deletion exposed another NLS which is masked in the full-length E2-TA protein. Control of nuclear entry by unmasking of NLSs has been described for NF-κB (14), the glucocorticoid receptor (31), and human heat shock factor 2 (36). The N-terminal NLS of the precursor form of NF-κB is blocked intramolecularly by a C-terminal domain, which is cleaved during maturation of the protein and allows the NLS to be exposed and NF-κB to localize to the nucleus. The mature form of NF-κB (p50-p65) is also regulated by masking of the NLS, but in this case masking is achieved by interaction with an inhibitor molecule, IκB (2). The human heat shock factor 2 (HSF2) protein contains two hydrophobic arms that interact to block the NLS. Under heat stress, a

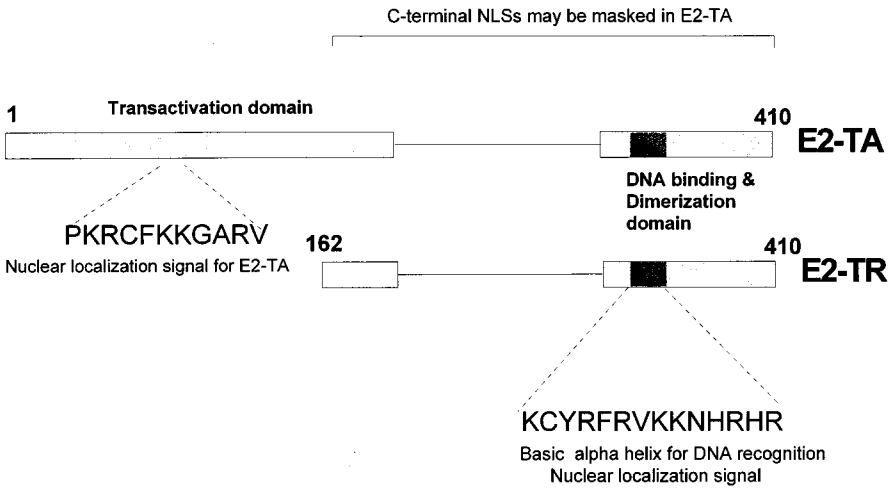

FIG. 7. Summary diagram of E2 NLSs.

conformational change takes place to expose the NLS, and HSF2 is taken into the nucleus. Studies are in progress to determine the mechanism by which the C-terminal NLSs are masked in the E2 protein.

A second NLS (BR3) located in the DNA-binding domain of E2 was identified by its ability to direct β-galactosidase to the nucleus. Mutation of this sequence in a functional protein consisting of the E2 DNA-binding domain disrupted nuclear localization and resulted in cells staining for E2 either exclusively in the cytoplasm, exclusively in the nucleus, or in both the nucleus and cytoplasm. In the background of the E2-TR, proteins with BR3 deletions or substitutions were found predominantly in the nucleus but a significant proportion were retained in the cytoplasm. It is difficult to determine whether the BR3 NLS is actually the predominant nuclear targeting signal in the E2-TR protein, because deletion of this region may allow other cryptic signals to be used that do not normally function in the wild-type protein. However, there are no other sequences that resemble typical NLSs in this polypeptide, and it is possible that the signal is not linear but depends on protein conformation. Alternatively, E2-TR proteins with BR3 mutations may enter the nucleus by other mechanisms. This analysis is also complicated by the fact that proteins with deletions or substitutions in BR3 are defective in dimerization and therefore are below the exclusion limits of the nuclear pores.

The X-ray crystallographic structure of the E2 DNA-binding domain bound to DNA shows that BR3 forms an alpha helix that contacts DNA (13). The basic residues within this helix are exposed on the surface of the protein, and some are crucial for recognition of specific base pairs in the E2 DNA-binding site. It is likely that the basic residues are also involved in BR3-mediated nuclear transport. Furthermore, in the tertiary structure of the E2 protein, a second, smaller alpha helix (aa 384 to 393), which contains conserved basic residues, lies in close proximity to BR3 (13). While this region does not appear to be important for DNA recognition, it could potentially form part of the NLS sequence in the E2 DNA-binding domain.

A number of other nuclear proteins such as NGFI-A (21), Fos (40), and Jun (6) contain NLSs in their DNA-binding domains, and it has been suggested that DNA-binding domains may have been archetypal targeting domains during the evolution of the nuclear membrane (9). In the case of Fos, the basic DNA recognition region can direct a cytoplasmic protein to the nucleus but deletion of this sequence did not disrupt nuclear localization of Fos (6), suggesting that Fos contains additional NLSs. The basic DNA recognition regions of the Fos/Jun family of proteins have some sequence similarity with the BR3 region of the E2 proteins, including a conserved cysteine residue that must be in the reduced state for DNA binding (1, 26). In the v-Jun protein, this cysteine has been mutated to serine, resulting in cell cycle-dependent nuclear localization (6). As shown in this study, an E2 protein containing a similar mutation in the redox-sensitive cysteine residue is nuclear. However, we are currently examining the localization of the E2 proteins at different stages of the cell cycle.

Nuclear localization of the yeast transcription factor SWI5 is regulated by CDC28 kinase-dependent phosphorylation at serines adjacent to the SWI5 NLS (29), and phosphorylation of serine residues near the NLS of SV40 large T-antigen by casein kinase II enhances nuclear import (17). The bovine papillomavirus type 1 E2 proteins are phosphorylated at two major sites near the basic region of the DNA-binding domain, and mutation of one of these sites results in greatly increased virus copy number (25). Studies are under way to determine whether E2 phosphorylation regulates nuclear transport.

A variety of nuclear proteins contain multiple NLSs; these include the p53 tumor suppressor, which contains several signals in its C-terminal end (35), and the MyoD protein, which contains two NLSs that can function independently (44). This study shows that the E2 proteins contain multiple NLSs. Moreover, although the repressor and transactivator proteins have overlapping sequences, they use different NLSs, and the C-terminal signals that are common to both proteins appear to be masked in the full-length E2-TA protein. This would suggest that at some time in the virus life cycle, these nuclear targeting signals could be differentially regulated. The ratio of the E2 transactivator and repressor proteins varies throughout the cell cycle: the transactivator is most abundant at S phase, when it is required for DNA replication, and the repressors are more abundant at other stages in the cell cycle (47). E2 expression is also regulated by cellular differentiation, and expression is limited to cells within the lower stratum spinosum in a papilloma (4). However, the latter studies could not discriminate between E2 transactivator and repressor species. The relative ratios and amounts of E2 protein present at specific times appear to be important in regulating the virus life cycle. Het-

erodimerization among E2 transactivator and repressor species may also be an important regulatory mechanism. Small changes in the concentration of E2 in the nuclear pool, as may occur when NLSs are masked, may be sufficient to significantly alter E2-mediated transactivation and viral DNA replication. Therefore, control of subcellular localization could be used to fine-tune the relative amounts of E2 transactivator and repressor proteins at certain points during the virus life cycle.

ACKNOWLEDGMENTS

We are very grateful to Jean-Michel Gauthier for the E2*dl*101–110, *dl*111–120, and *dl*41–50 plasmids; to Patricia Winokur for generating several of the PAVA E2 viruses; to Dan DiMaio for the PAVA E2-TR virus; and to Elliot Androphy for the B201 and B202 antibodies. We thank Scott VandePol, John Hanover, and Jon Yewdell for critical reading of the manuscript.

REFERENCES

1. **Abate, C., L. Patel, F. J. Rauscher III., and T. Curran.** 1990. Redox regulation of Fos and Jun DNA-binding activity in vitro. Science **249:**1157–1161.
2. **Baeuerle, P. A., and D. Baltimore.** 1988. Activation of DNA binding activity in an apparently cytoplasmic precursor of the NF-κB transcription factor. Cell **53:**211–217.
3. **Barsoum, J., S. S. Prakash, P. Han, and E. J. Androphy.** 1992. Mechanism of action of the papillomavirus E2 repressor: repression in the absence of DNA binding. J. Virol. **66:**3941–3945.
4. **Burnett, S., A.-C. Ström, N. Jareborg, A. Alderborn, J. Dillner, J. Moreno-Lopez, U. Pettersson, and U. Kiessling.** 1990. Induction of bovine papillomavirus E2 gene expression and early region transcription by cell growth arrest: correlation with viral DNA amplification and evidence for differential promoter induction. J. Virol. **64:**5529–5541.
5. **Carruth, M., and A. A. McBride.** 1995. Unpublished data.
6. **Chida, K., and P. Vogt.** 1992. Nuclear translocation of viral Jun but not of cellular Jun is cell cycle dependent. Proc. Natl. Acad. Sci. USA **89:**4290–4294.
7. **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.
8. **Coleman, N. A., and M. E. Peeples.** 1993. The matrix protein of Newcastle disease virus localizes to the nucleus via a bipartite nuclear localization signal. Virology **195:**596–607.
9. **Dingwall, C., and R. A. Laskey.** 1991. Nuclear targeting sequences—a consensus? Trends Biochem. Sci. **16:**478–481.
10. **Fabre, E., and E. C. Hurt.** 1994. Nuclear transport. Curr. Opin. Cell Biol. **6:**335–342.
11. **Gerard, R. D., and Y. Gluzman.** 1985. New host cell system for regulated simian virus 40 DNA replication. Mol. Cell. Biol. **5:**3231–3240.
12. **Ham, J., N. Dostatni, J. M. Gauthier, and M. Yaniv.** 1991. The papillomavirus E2 protein: a factor with many talents. Trends Biochem. Sci. **16:**440–444.
13. **Hegde, R. S., S. R. Grossman, L. A. Laimins, and P. B. Sigler.** 1992. Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. Nature (London) **359:**505–512.
14. **Henkel, T., U. Zabel, K. Van Zee, J. M. Muller, E. Fanning, and P. A. Baeuerle.** 1992. Intramolecular masking of the nuclear location signal and dimerization domain in the precursor for the p50 NF-κB subunit. Cell **68:**1121–1133.
15. **Howley, P. M.** 1995. Papillomaviruses and their replication, p. 2045–2076. *In* B. Fields (ed.), Virology, 3rd ed. Raven Press, New York.
16. **Hubbert, N. L., J. T. Schiller, D. R. Lowy, and E. J. Androphy.** 1988. Bovine papilloma virus-transformed cells contain multiple E2 proteins. Proc. Natl. Acad. Sci. USA **85:**5864–5868.
17. **Jans, D. A., and P. Jans.** 1995. Negative charge at the casein kinase II site flanking the nuclear localization signal of the SV40 large T-antigen is mechanistically important for enhanced nuclear import. Oncogene **9:**2961–2968.
18. **Lambert, P. F., B. A. Spalholz, and P. M. Howley.** 1987. A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 transactivator. Cell **50:**69–78.
19. **Li, R. and M. R. Botchan.** 1993. The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate *in vitro* BPV-1 DNA replication. Cell **73:**1207–1221.
20. **Li, R., and M. R. Botchan.** 1994. Acidic transcription factors alleviate nucleosome-mediated repression of DNA replication of bovine papillomavirus type 1. Proc. Natl. Acad. Sci. USA **91:**7051–7055.
21. **Matheny, C., M. L. Day, and J. Milbrandt.** 1994. The nuclear localization signal of NGFI-A is located within the zinc finger DNA binding domain. J. Biol. Chem. **269:**8176–8181.
22. **McBride, A. A.** 1995. Unpublished observations.
23. **McBride, A. A., J. B. Bolen, and P. M. Howley.** 1989. Phosphorylation sites of the E2 transcriptional regulatory proteins of bovine papillomavirus type 1. J. Virol. **63:**5076–5085.
24. **McBride, A. A., J. C. Byrne, and P. M. Howley.** 1989. E2 polypeptides encoded by bovine papillomavirus type 1 form dimers through the common carboxyl-terminal domain: transactivation is mediated by the conserved amino-terminal domain. Proc. Natl. Acad. Sci. USA **86:**510–514.
25. **McBride, A. A., and P. M. Howley.** 1991. Bovine papillomavirus with a mutation in the E2 serine 301 phosphorylation site replicates at a high copy number. J. Virol. **65:**6528–6534.
26. **McBride, A. A., R. D. Klausner, and P. M. Howley.** 1992. Conserved cysteine residue in the DNA-binding domain of the bovine papillomavirus type 1 E2 protein confers redox regulation of the DNA-binding activity in vitro. Proc. Natl. Acad. Sci. USA **89:**7531–7535.
27. **McBride, A. A., H. R. Romanczuk, and P. M. Howley.** 1991. The papillomavirus E2 regulatory proteins. J. Biol. Chem. **266:**18411–18414.
28. **McBride, A. A., R. Schlegel, and P. M. Howley.** 1988. The carboxy-terminal domain shared by the bovine papillomavirus E2 transactivator and repressor proteins contains a specific DNA binding activity. EMBO J. **7:**533–539.
29. **Moll, T., G. Tebb, U. Surana, H. Robitsch, and K. Nasmyth.** 1991. The role of phosphorylation and the CDC28 protein kinase in the cell cycle-regulated nuclear import of the S. cerevisiae transcription factor SW15. Cell **66:**743–758.
30. **Moreland, R. B., G. L. Langevin, R. H. Singer, R. L. Garcea, and L. M. Hereford.** 1987. Amino acid sequences that determine the nuclear localization of yeast histone 2B. Mol. Cell. Biol. **7:**4048–4057.
31. **Picard, D., and K. R. Yamamoto.** 1987. Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. EMBO J. **6:**3333–3340.
32. **Robbins, J., S. Dilworth, R. A. Laskey, and C. Dingwall.** 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell **64:**615–623.
33. **Seo, Y., F. Muller, M. Lusky, E. Gibbs, H. Kim, B. Phillips, and J. Hurwitz.** 1993. Bovine papilloma virus (BPV)-encoded E2 protein enhances binding of E1 protein to the BPV replication origin. Proc. Natl. Acad. Sci. USA **90:**2865–2869.
34. **Settleman, J., and D. DiMaio.** 1988. Efficient transactivation and morphologic transformation by bovine papillomavirus genes expressed from a bovine papillomavirus/simian virus 40 recombinant virus. Proc. Natl. Acad. Sci. USA **85:**9007–9011.
35. **Shaulsky, G., N. Goldfinger, A. Ben-Ze'ev, and V. Rotter.** 1990. Nuclear accumulation of p53 protein is mediated by several nuclear localization signals and plays a role in tumorigenesis. Mol. Cell. Biol. **10:**6565–6577.
36. **Sheldon, L. A., and R. E. Kingston.** 1993. Hydrophobic coiled-coil domains regulate the subcellular localization of human heat shock factor 2. Genes Dev. **7:**1549–1558.
37. **Silver, P. A.** 1991. How proteins enter the nucleus. Cell **64:**489–497.
38. **Skiadopoulos, M. H., and A. A. McBride.** 1995. Unpublished observations.
39. **Spalholz, B. A., A. A. McBride, T. Sarafi, and J. Quintero.** 1993. Binding of bovine papillomavirus E1 to the origin is not sufficient for DNA replication. Virology **193:**201–212.
40. **Tratner, I., and I. M. Verma.** 1991. Identification of a nuclear targeting sequence in the fos protein. Oncogene **6:**2049–2053.
41. **Ustav, E., M. Ustav, P. Szymanski, and A. Stenlund.** 1993. The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. Proc. Natl. Acad. Sci. USA **90:**898–902.
42. **Ustav, M., and A. Stenlund.** 1991. Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. EMBO J. **10:**449–457.
43. **Ustav, M., E. Ustav, P. Szymanski, and A. Stenlund.** 1991. Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. EMBO J. **10:**4321–4329.
44. **Vandromme, M., J. Cavadore, A. Bonnieu, A. Froeschle, N. Lamb, and A. Fernandez.** 1995. Two nuclear localization signals present in the basic-helix 1 domains of myoD promote its active nuclear translocation and can function independently. Proc. Natl. Acad. Sci. USA **92:**4646–4650.
45. **Winokur, P. L., and A. A. McBride.** Unpublished data.
46. **Yang, L., R. Li, I. J. Mohr, R. Clark, and M. R. Botchan.** 1991. Activation of BPV-1 replication *in vitro* by the transcription factor E2. Nature (London) **353:**628–632.
47. **Yang, L., I. Mohr, R. Li, T. Nottoli, S. Sun, and M. Botchan.** 1991. Transcription factor E2 regulates BPV-1 DNA replication *in vitro* by direct protein-protein interaction. Cold Spring Harbor Symp. Quant. Biol. **56:**335–346.